

**TRANSACTIONS
OF THE
AMERICAN MATHEMATICAL SOCIETY**

EDITED BY

WILLIAM C. GRAUSTEIN

EINAR HILLE

C. C. MACDUFFEE

WITH THE COÖPERATION OF

A. A. ALBERT

T. H. HILDEBRANDT

MARSTON MORSE

H. P. ROBERTSON

GABOR SZEGÖ

ERIC T. BELL

E. P. LANE

OYSTEIN ORE

M. H. STONE

G. T. WHYBURN

JESSE DOUGLAS

R. E. LANGER

H. L. RIETZ

J. L. SYNGE

OSCAR ZARISKI

VOLUME 46

JULY TO DECEMBER, 1939

PUBLISHED BY THE SOCIETY

MENASHA, WIS., AND NEW YORK

1939

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

Chester C. Corbin Library Fund
Bound June, 1940

Composed, Printed and Bound by
~~The College Press~~
George Banta Publishing Company
Menasha, Wisconsin

63660

QA
1
F001
v.46

TABLE OF CONTENTS

VOLUME 46, JULY TO DECEMBER, 1939

ADAMS, C. R., and CLARKSON, J. A. A correction to "Properties of functions $f(x, y)$ of bounded variation".....	468
BAER, R. Nets and groups.....	110
BELL, P. O. A study of curved surfaces by means of certain associated ruled surfaces.....	389
BOAS, R. P. On a generalization of the Stieltjes moment problem.....	142
CAMERON, R. H., and WIENER, N. Convergence properties of analytic functions of Fourier-Stieltjes transforms.....	97
CLARKSON, J. A., and ADAMS, C. R. A correction to "Properties of functions $f(x, y)$ of bounded variation".....	468
CRAMÉR, H. On the representation of a function by certain Fourier integrals.....	191
DE CICCIO, J. The differential geometry of series of lineal elements.....	348
DILWORTH, R. P. Non-commutative residuated lattices.....	426
GREVILLE, T. N. E. Invariance of the admissibility of numbers under certain general types of transformations.....	410
HARTMAN, P. Mean motions and almost periodic functions.....	66
LANGER, R. E. The boundary problem of an ordinary linear differential system in the complex domain.....	151
LANGER, R. E. A correction to "The boundary problem of an ordinary linear differential system in the complex domain".....	467
LEHMER, D. H. On the remainders and convergence of the series for the partition function.....	362
LEWIS, D. C. Contributions to the transformation theory of dynamics... ..	374
MACCOLL, L. A. Geometric aspects of relativistic dynamics.....	328
MAC LANE, S. Steinitz field towers for modular fields.....	23
OLDENBURGER, R. Exponent trajectories in symbolic dynamics.....	453
PERLIS, S. Maximal orders in rational cyclic algebras of composite degree..	82

RAUDENBUSH, H. W., and RITT, J. F. Ideal theory and algebraic difference equations.	445
RINEHART, R. F. An interpretation of the index of inertia of the discriminant matrices of a linear associative algebra.	307
RITT, J. F., and RAUDENBUSH, H. W. Ideal theory and algebraic difference equations.	445
TRJITZINSKY, W. J. General theory of singular integral equations with real kernels.	202
WALD, A. Limits of a distribution function determined by absolute moments and inequalities satisfied by absolute moments.	280
WALSH, J. L. On interpolation by functions analytic and bounded in a given region.	46
WIENER, N., and CAMERON, R. H. Convergence properties of analytic functions of Fourier-Stieltjes transforms.	97
ZORN, M. Continuous groups and Schwarz' lemma.	1

CONTINUOUS GROUPS AND SCHWARZ' LEMMA*

BY

MAX ZORN

INTRODUCTION

The famous lemma of H. A. Schwarz is doubtless one of the basic theorems in the theory of analytic functions. In this paper I propose to study the lemma from a topological point of view. The results have been announced, without proof, in a previous note.† Several changes, corrections, and additions have been made; I use the opportunity to state here my indebtedness to D. W. Hall for his inspiring interest and helpful criticism.

The theory to be presented is a by-product of a more comprehensive treatment of conformal mappings‡ which will be communicated elsewhere.

Like the theories of Kerékjártó§ and Stoilow|| our investigations are made with a direct view to the characterization of conformal mappings. Yet both authors deal with the conformal mappings individually, whereas we aim more at the characterization of the system of all conformal mappings of a Riemann surface S in itself. As an equivalent to this simplification of the problem we attempt to keep the space S general as long as possible, whereas usually S is supposed to be locally plane from the outset.

The theory of Schwarz' lemma has been separated from the rest because of its independence and also because it seems to be of value for the study of the hardest characterization problem, the problem of Brouwer.

The present paper is divided into three parts. Part I is of a rather general nature and can be read without any topological preparation. For the other parts a certain familiarity with topological notions and theorems is necessary. Parts I and II together lead up to a theorem which is formally identical with the Schwarz lemma. In III, particularly in §8, we show that this formal identity is material identity; in §9 we derive, with the aid of the geometric theorems from II some interesting topological features of the underlying space.

* Presented to the Society, February 25, 1939; received by the editors September 9, 1938.

† *Sur le lemme de Schwarz*, Comptes Rendus de l'Académie des Sciences, vol. 206, p. 725.

‡ Cf. *Topological studies in the theory of analytic functions*, Bulletin of the American Mathematical Society, abstract 43-11-415.

§ Cf. B. de Kerékjártó, *Sur la structure des transformations topologiques . . .*, Enseignement Mathématique, vol. 35 (1936), p. 297.

|| Stoilow, *Leçons sur les Principes Topologiques de la Théorie des Fonctions Analytiques*, Paris, 1938.

Notations. We use only italic letters; consequently, concepts of different logical types are often denoted by letters of the same alphabet: d, i, m, n, a are indices, d, i and m, n natural numbers, a is arbitrary; e, p, q, r, s, x, y, z , are points (most of them in S); $S, A, C, E, K, L, O, U, V$ denote sets of points, usually contained in S ; $F, F^*, F', G, H, P, R, R_p, R_i, R^{(x)}$ are transformations, usually continuous single-valued mappings of S in itself; N is a family of transformations; in general the transformations F , and so on, will belong to N .

If all x_i are in a set such as C , we call x_i a sequence from C . A "subsequence" of a sequence, say x_i , is formed by choosing an increasing sequence ($i_{n+1} > i_n$) of indices; it is convenient to denote the new sequence by x'_i ; a subsequence of the subsequence would be written x''_i .

Theorems and definitions are numbered together; a definition is indicated by brackets, a theorem by parentheses.

PART I

1. Continuous transformations. We make the following definition:

[1.1] S is a topological space which we assume to be metrizable. The metric of the space does not occur explicitly, but we shall have to use limit relations like $\lim x_n = x$, functions like the closure \bar{A} , the boundary $\text{Bd}(A)$, the frontier $\text{Fr}(A)$, and properties like open, closed, connected, locally connected; the terms compact, limited are defined explicitly for obvious reasons.

In Part I, however, we do not need all the consequences of the metrizability; it is sufficient to assume that S is an L^* -space as defined, for example, in Kuratowski's book on topology.†

[1.1.1] S is an L^* -space if convergence of sequences is defined and satisfies the following conditions:

- I. If $\lim x_n = x$, then $\lim x'_n = x$.
- II. If $x_n = x$, then $\lim x_n = x$.
- III. If for every subsequence x'_n of x_n a subsequence x''_n with $\lim x''_n = x$ can be found, then $\lim x_n = x$.

[1.2] A point x is a limit point of A if a sequence x_i from A exists such that $\lim x_i = x$. A point x is a limit point of a sequence x_i if a subsequence x'_i with $\lim x'_i = x$ exists.

If every sequence x_i from A has at least one limit point, then A is called "limited."

A is "compact" if every sequence from A has a limit point in A .

† C. Kuratowski, *Topologie I*, Warsaw, 1933, pp. 76-77; cf. also the literature mentioned there.

In the sequel we shall be concerned mostly with a family N of single-valued, continuous transformations F, G, H, I, P, R, \dots . The domain (of definition) is always S , the range (of values) $F(S)$ is a subset of S . The natural definition of continuity in L^* -spaces is the following:

[1.3] F is continuous if $\lim x_n = x$ implies $\lim F(x_n) = F(x)$.

Convergent sequences of (continuous) functions F will occur rather often; it seems that the type of convergence which has been introduced as "continuous convergence"† is the most appropriate one for the abstract theory of conformal mappings.

[1.4] A sequence of transformations F_n is said to converge towards F , that is, $\lim F_n = F$, if $\lim x_n = x$ implies $\lim F_n(x_n) = F(x)$.

Obviously this implies $\lim F_n(x) = F(x)$; but the converse is not true.

By virtue of the definition [1.4] any set of continuous mappings F forms an L^* -space. We shall have to use the corresponding property III in our proofs; hence we state explicitly the following theorem:

(1.4.1) If every subsequence F_n' contains a subsequence F_n'' with $\lim F_n'' = F$, then $\lim F_n = F$.

Indeed, let $\lim x_n = x$, and consider the sequence of points $F_n(x_n)$. From every subsequence $F_n'(x_n')$ we can select $F_n''(x_n'')$ such that $\lim F_n''(x_n'') = F(x)$. Consequently, $\lim F_n(x_n) = F(x)$, which implies $\lim F_n = F$.

[1.5] A sequence F_n is called "properly divergent" if for no point x the sequence $F_n(x)$ has a limit point.

We recall the usual notations and conventions about composition of functions:

[1.6] The product $H = FG$ of F and G (in this order) is defined by $H(x) = FG(x) = F(G(x))$. The identity is the transformation I which leaves all points invariant, $I(x) = x$. A function G is the inverse of F if $GF = I$; it may not exist, but if it does, then it is unique and satisfies $FG = I$. Powers F^n are defined as usual; if the inverse exists, it is always written as F^{-1} .

(1.7) If $\lim F_n = F$ and $\lim G_n = G$, then $\lim F_n G_n = FG$.

This is an immediate consequence of the "continuity" of the convergence.

Let $\lim x_n = x$. It follows from $\lim G_n(x_n) = G(x)$ that $\lim F_n(G_n(x_n)) = F(\lim G_n(x_n)) = F(G(x)) = FG(x)$.

(1.8) If $\lim F_n = F$ and $\lim F_n^{-1} = G$, then $G = F^{-1}$.

In other words, if the inverses F_n^{-1} exist and converge towards a limit, then

† Cf. C. Carathéodory, *Conformal Representation*, Cambridge Tracts, no. 28.

$$\lim F_n^{-1} = (\lim F_n)^{-1}.$$

The proof is an algebraic consequence of (1.7), for $GF = (\lim F_n^{-1})(\lim F_n) = \lim F_n^{-1}F_n = I$.

(1.9) *F is called nilpotent if a point p exists such that $\lim x_n = x$ implies*

$$\lim F^n(x_n) = p.$$

[1.10] *The transformation which maps every point on the same point p is called "constant" and denoted by P.*

With this terminology we can say that *F* is nilpotent exactly if its powers converge towards a constant.

(1.11) *The point p is a fixed point of F, $F(p) = p$. It is also the only fixed point of F.*

Indeed, writing *p* in the form $\lim F^n(p)$, we obtain $F(p) = F(\lim F^n(p)) = \lim F^{n+1}(p) = \lim F^n(p) = p$.

If, on the other hand, $F(q) = q$, the relations $F^n(q) = q$, $\lim F^n(q) = q$, and $\lim F^n(q) = p$ give $q = p$.

2. **The family *N*.** In this section we introduce a group of definitions and assumptions which describe abstractly some features of the analytic mappings of the unit circle in itself.

[2.1] *The family N is a set of transformations with the following properties:*

I. Continuity. *The elements of N are single-valued continuous transformations of S into itself, $F(S) \subset S$.*

II. Composition, identity. *The identity I is in N, and if F and G are in N, then their product FG is in N.*

III. Cancellation. *If F, G, H are in N, and if F is not constant, then the equality $GF = HF$ implies $G = H$.*

IV. Normality. *Every sequence F_n from N contains a subsequence $F_{n'}$ which is either properly divergent or else converges towards an element F of N.*

If we want *S* to be the unit circle of the complex number plane and *N* the set of all analytic mappings $F, F(S) \subset S$, we speak of "the classical case."

In the classical case, I-IV are fulfilled; I-III are elementary, whereas IV has perhaps a more advanced character and belongs to the theory of normal families.[†]

From these assumptions alone we shall derive a topological version of the Schwarz lemma. In Part II a geometrical formulation will be established on

[†] Cf. R. Montel, *Sur les Familles Normales de Fonctions Analytiques et leurs Applications*, Paris, Gauthier-Villars, 1927; cf. also Kerékjártó, loc. cit., p. 308; K. Szilárd, *Untersuchungen ueber die Grundlagen der Funktionentheorie*, Mathematische Zeitschrift, vol. 26, p. 653.

the basis of further restrictions on S and N . Finally we show how the abstract theorem yields the ordinary Schwarz lemma in the classical case.

The geometry in S will be provided by those elements of N which have an inverse in N . In particular, the analogue of the ordinary rotations is of use.

[2.2] *A transformation R is called a rotation if*

- (a) *R is in N ;*
- (b) *the inverse R^{-1} exists;*
- (c) *R^{-1} is in N ;*
- (d) *R has a fixed point p .*

We shall also say that R is a rotation "about p " or "with center p "; the fixed point will often be indicated by the subscript p : $R_p(p) = p$.

The point p , unless stated otherwise, may be considered fixed in advance. In particular, it will be fixed for the following definitions of "rotatory," "invariant," "circumference."

(2.3) *The rotations about p form a group.*

That means that $R_1 R_2$ is a rotation, $(R_1 R_2) R_3 = R_1 (R_2 R_3)$, I is a rotation, and R_1^{-1} is a rotation satisfying $R_1^{-1} R_1 = R_1 R_1^{-1} = I$. (The proof is omitted.)

[2.4] *If the set A contains its image $R_p(A)$ under every rotation, it is called invariant. Since R_p^{-1} is also a rotation, we might have said $R_p(A) = A$.*

[2.5] *A set A is "rotatory" if for any two points q, r in A , there exists a rotation R_p such that $R_p(q) = r$.*

[2.6] *A set which contains a point q is called a circumference L_q if it is invariant and rotatory.*

If necessary, we say "circumference through q with center p ."

This definition is justified by the fact that L_q consists of all points of the form $R_p(q)$.

The definitions and assumptions set forward in these two introductory paragraphs enable us now to formulate and prove the first (topological) version of Schwarz' lemma. The rotations will hereby play a quite important role; we shall establish first some of their properties.

3. Rotations and circumferences. We make the following assertion:

(3.1) *Let $\{F_a\}$ be a subset of N , the index a ranging over an arbitrary set of symbols. If for one single point p the set $\{F_a(p)\}$ is limited, then for every point x the set $\{F_a(x)\}$ is limited.*

We derive this from the normality property IV in the following more general form:

(3.1.1) *If $\{F_\alpha(p)\}$ is limited, then every sequence F_{a_i} contains a subsequence $F_{a_i'}$ which converges towards an element of N .*

Indeed, we only have to select a subsequence $F_{a_i'}$ which is either convergent or properly divergent. The second possibility cannot arise, since $F_{a_i'}(p)$ has at least one limit point. Consequently, any sequence $F_{a_i}(x)$ has at least the limit point $\lim F_{a_i'}(x)$. From the theorem (3.1) we shall generally use the following special case:

(3.1.2) *If $F_i(p) = p$, then F_i has a convergent subsequence F_i' .*

Two other consequences are the following:

(3.1.3) *The circumferences L_q are limited.*

(3.1.4) *If the transformations F_i are in N , and if the sequence F_i converges "pointwise" towards F , that is, for every x $\lim F_i(x) = F(x)$, then F is in N , and the convergence*

$$\lim F_i = F$$

is continuous.

The theorem (3.1.3) is obvious since L_q consists of all points $R_p(q)$, and $R_p(p) = p$. We shall afterwards show that L_q is even compact. The second statement is based on (1.4.1), and we prove it in a more general form. We do not need the generalization; it is inserted merely as the abstract background of the theorems of Stieltjes-Porter-Vitali-Blaschke.†

(3.1.5) *Let A be such that $F(x) = G(x)$ for all x in A implies that $F = G$ in S , in case F and G are in N . Suppose that for all x in A $\lim F_i(x)$ exists. Then $F_i(x)$ converges for all x in S , towards say $F(x)$; $F(x)$ is in N , and we have $\lim F_i = F$.*

Indeed, for every subsequence F_n' there exists a convergent subsequence F_n'' , with the limit F'' contained in N but formally dependent on the subsequence F_n'' . Yet all these possible limit functions are identical on A ; consequently, they are identical throughout. That is sufficient (cf. (1.4.1)) for the relation $\lim F_i = F$.

The foregoing theorems are now applied in the case of rotations (about p).

(3.2) *Every sequence R_n of rotations contains a convergent subsequence; the limit mapping is in N .*

This is again a special case of (3.1.2). But we can make the following stronger statement:

(3.3) *The limit of a sequence R_i of rotations is again a rotation R .*

† Cf. Bieberbach, *Funktionentheorie*, vol. 2, 1st edition, p. 158.

Anticipating the result, we write $\lim R_i = R$. Since $R_i(p) = p$ implies $R(p) = p$, and R is (cf. (3.2)) in N , we have only to prove that R has an inverse in N .

Consider the sequence of rotations R_i^{-1} . There will be at least one convergent subsequence $R_{i'}^{-1}$, $\lim R_{i'}^{-1} = G$, where G is in N . Since the limit of the corresponding sequence $R_{i'}$ is R , (1.8) yields that G is the inverse R^{-1} of R .

(3.4) *If $\lim R_i = R$, then $\lim R_i^{-1} = R^{-1}$.*

Take any subsequence $R_{i'}^{-1}$ of the sequence R_i^{-1} . The proof of the foregoing statement shows that we can select a convergent subsequence $R_{i''}^{-1}$ which converges towards R^{-1} . On account of (1.4.1) this implies $\lim R_{i'}^{-1} = R^{-1}$.

These theorems may be condensed into the statement that the rotations about p , under the continuous convergence, form a compact L^* -group.

(3.5) *The circumferences L_q are compact.*

Let q_i be a sequence from L_q ; then by definition $q_i = R_i(q)$. Selecting a convergent subsequence $R_{i'}$ with the limit R' we see that the corresponding subsequence $q_{i'} = R_{i'}(q)$ has the limit point $R'(q)$, which is in L_q .

(3.6) *If S has more than one point, then a rotation is not nilpotent.*

4. Topological version of the lemma. The theorem (4.1) is, in the classical case, one of the numerous consequences of Schwarz' lemma. It expresses as far as possible the tendency of a mapping F which has a fixed point p but is not a rotation, to move the points of S "nearer" p . Why we call it a topological version of the classical lemma will be evident afterwards, when the application to the classical case is made.

(4.1) *A transformation F in N with the fixed point p is either a rotation (about p) or is nilpotent.*

The proof is made in two steps, (4.2) and (4.3). We show first that F is already nilpotent if only one subsequence F^{n_i} of the sequence F^n converges towards the constant P . If then F is not nilpotent, there must be a convergent sequence F^{m_i} with a nonconstant limit F^* in N . It is shown in (4.3) that in this case F has an inverse F^{-1} in N , which is more or less explicitly constructed as a limit of a sequence of powers of F .

All transformations occurring in this paragraph will be in N , either by assumption (as in the case of F) or because they are limits of mappings in N .

(4.2) *If $F(p) = p$ and if a sequence F^{n_i} , where $n_{i+1} > n_i$, tends towards the constant P , then F is nilpotent and $\lim F^n = P$.*

It is sufficient to show that every sequence F^{m_i} , $m_{i+1} > m_i$, contains a convergent subsequence $F^{m'_i}$ with the limit P .

We select two subsequences m'_i, n'_i such that

- (a) $d_i = n'_i - m'_i$ is increasing, and
- (b) the sequence F^{d_i} is convergent with $\lim F^{d_i} = F'$.

Such a sequence exists; the first condition can be fulfilled because n_i and m_i are strictly increasing; the second condition, because $F(p) = p$, $F^n(p) = p$.

The relations $\lim F^{n'_i} = \lim F^{n_i} = P$, $\lim F^{d_i} = F'$ imply (cf. (1.7)) that $\lim F^{m'_i} = \lim F^{n'_i} F^{d_i} = P F' = P$.

We note that while the normality has been used freely, the cancellation property has not yet appeared in the proofs.

(4.3) *If $F(p) = p$ and if a convergent sequence F^{n_i} with the nonconstant limit F^* exists, then F has an inverse in N .*

The proof is somewhat similar to the preceding one, but F' is defined slightly differently and the cancellation property III is essential.

We select again a subsequence n'_i such that

- (a) $d_i = n'_{i+1} - n'_i - 1$ is strictly increasing, and
- (b) F^{d_i} is convergent with $\lim F^{d_i} = F'$.

From these assumptions we derive the equality $F^* = F^* F' F$, for

$$\lim F^{n_i} = \lim F^{n'_{i+1}} = \lim F^{n'_i} F^{d_i} F = (\lim F^{n'_i})(\lim F^{d_i}) F.$$

Writing this in the form $F^* I = F^* (F' F)$ and using the fact that F^* is not a constant, we obtain, by virtue of the cancellation law, $F' F = I$.

In F' we have, therefore, the inverse of F , the existence of which was asserted in our theorem.

The principal theorem follows now as indicated before. If $F(p) = p$, then $F^n(p) = p$ shows that a convergent sequence F^{n_i} exists. If the limit is constant, then (4.2) implies that F is nilpotent; if it is not, then (4.3) shows that F has an inverse and consequently is a rotation (about p).

As an interesting corollary we obtain the following:

(4.4) *A transformation with two different fixed points p and q is a rotation.*

In the classical case one knows more: the rotation F is the identity. This generalization suggests itself as an additional axiom, which (see the end of the paper) permits a more precise description of S and N . In view of the theory of Kerékjártó we call attention to the fact that instead of N we could have studied the subsystem formed by powers of F and their limits.

PART II

5. New restrictions on S and N . From now on we shall use more freely the topological terminology, indicated by the words open, neighborhood, closed; closure \bar{A} , boundary $\text{Bd}(A)$ of a set A ; connected, component, locally con-

nected; separate, cut point; semicompact, locally compact, (perfectly) separable, and metrizable.

The set AB is the common part and $A+B$ the union of the two sets, and $S-A$ is the complement of A in S .

[5.1] S is now a metrizable space with the following additional properties:

- I. S is connected and contains more than one point.
- II. S is locally connected.
- III. S has no cut points; that is, for all points x the set $S - \{x\}$ is connected.

In Part II we shall use III generally for $x=p$, where p is arbitrary but fixed.

Before we set down the restrictions on N we define the geometrical concepts "circle" and "closed circle."

[5.2] The component of $S - L_q$ which contains p is called the circle with center p determined by q and is denoted by C_q .

In other words, the circle is the largest connected subset of the complement of the circumference L_q which contains p . If (and only if) q is identical with p , then C_q is empty.

We note that this describes the interior of the circular area determined by a circular curve in euclidean geometry, which has p as center and q on the curve.

[5.2.1] The closure \bar{C}_q of C_q , comprising C_q and all its limit points, is called a closed circle.

The restrictions on N are now phrased as properties of circles and circumferences.

[5.3] N is from now on a family of transformations which has not only the properties I-IV of [2.1] but the following:

V. If $q \neq p$, then L_q separates S ; that is, $S - L_q$ is not connected.

VI. The space S is not representable as a finite sum of circles (with possibly different centers).†

These two axioms constitute very heavy restrictions on N , but in exchange we obtain a quite rich geometry (topologically speaking) for S .

6. **Circles and circumferences.** We can make the following assertion:

(6.1) The circles C_q are open and connected.

† This property was not contained in the before mentioned note; my proof for the central theorem, loc. cit. (II, 5), contained a mistake which was pointed out to me by Mr. Hall and which I was not able to correct without a new assumption. The particular form of VI has been chosen since it is also useful for the justification theorem (§8).

If $q = p$, then $L_q = \{p\}$, and $S - L_q$ does not contain p . In this case we have to interpret C_q as the empty set, which may be considered open and connected.

If $q \neq p$, then p is in $S - L_q$, since $R(q) = p$ implies $q = R^{-1}(p) = p$.

L_q is closed (even compact); its complement $S - L_q$ is consequently open. Now S is locally connected; that is, in every neighborhood U_x (open set containing x) there exists a neighborhood V_x which is connected.

It follows that a component (largest connected subset) of any open set in a locally connected space is open; C_q is such a component, hence it is open; it is connected by definition. If it is not empty, it contains p .

(6.2) *The circles C_q and their boundaries $\text{Bd}(C_q)$ are invariant.*

This is a consequence of the following group of statements:

$$(6.2.1) \quad R(A+B) = R(A) + R(B); R(AB) = R(A)R(B); R(S-A) = S - R(A).$$

This holds for subsets A, B of S and for any (1-1) mapping of S on itself, in particular, for a rotation.

$$(6.2.2) \quad R(\overline{A}) = \overline{R(A)}.$$

This holds at least for topological mappings (where R and R^{-1} are continuous).

$$(6.2.3) \quad R(\text{Bd}(A)) = \text{Bd}(R(A)).$$

The boundary, as the set of all points which are limit points of sequences from A but not in A , can be written as

$$\text{Bd}(A) = \overline{A} - A.$$

(6.2.3) follows algebraically from this definition and the preceding identities.

(6.2.4) *Any function of invariant sets A, B, C which is composed from sums, products, complements, and closures is invariant.*

For example the "frontier $\text{Fr}(A)$ of A " is equal to $R(\text{Fr}(A))$ because by definition

$$\begin{aligned} \text{Fr}(A) &= \overline{A(S-A)}, & R(\overline{A \cdot S - A}) &= R(\overline{A})R(\overline{S-A}), \\ R(\overline{A}) &= \overline{R(A)} = \overline{A}, & R(\overline{S-A}) &= \overline{R(S-A)} = \overline{R(S) - R(A)} = \overline{S-A}. \end{aligned}$$

Also

$$R(\text{Bd}(A)) = R(\overline{A} - A) = R(\overline{A}) - R(A) = \overline{R(A)} - R(A) = \text{Bd}(R(A)).$$

In order to derive (6.2) we have only to go back to the definition of C_q . The set L_q is invariant; hence $S - L_q$ is invariant; a rotation R maps $S - L_q$

topologically on itself, a connected subset on a connected subset, a largest connected subset on a (possibly different) subset of the same character, and, since $R(p) = p$, the component C_q of p on itself.

The boundary $\text{Bd}(C_q)$ is invariant as a function of an invariant set; this invariance we use now for the determination of $\text{Bd}(C_q)$.

(6.3) *The boundary $\text{Bd}(C_q)$ is exactly L_q , if $q \neq p$; if $q = p$ it is, of course, empty.*

If $q = p$, then $C_q = 0$; hence $\text{Fr}(C_q) = \text{Fr}(0) = 0$. Hence we assume $C_q \neq 0$. Since q is in L_q , q is not in C_q and C_q is not equal to S .

The set C_q could not be closed, for an open and closed set in a connected space S is either 0 or S . Hence there is a point which is limit point for C_q but not in C_q ; let r be such a boundary point. The point r cannot be in $S - L_q$, for C_q is a component of $S - L_q$; hence it contains all its limit points in $S - L_q$, and it is "relatively closed" with respect to $S - L_q$. The point r , that is, any boundary point of C_q , is therefore in L_q .

The boundary is not only a non-empty subset of L_q , it is also invariant. Since 0 and L_q are the only invariant subsets of L_q , the boundary C_q is exactly L_q .

The connectedness of S and C will be used so often in the proofs to come that we deem it advisable to insert the following theorem:

(6.3.1) *The connectedness of a space is equivalent to the following implications:*

- (a) *If a set A is open and closed, it is either 0 or the whole space.*
- (b) *If an open set A has no boundary, then it is 0 or the whole space.*
- (c) *If one knows, for an open set A , that $\text{Bd}(A) \subset A$, then A is 0 or the whole space.*
- (d) *The space is not the sum of two disjoint open proper subsets.*

These statements are trivial consequences of the following definition:

[6.3.2] *A space S is connected if $A + B = S$ and $\overline{AB} = 0$ imply that either \overline{A} or \overline{B} is empty.*

Since $\text{Bd}(A) = \overline{A} - A$, we get from (6.3) the corollary:

(6.3.3) $\overline{C}_q = C_q + L_q$, if $p \neq q$.

For $p = q$ this is not true since $\overline{C}_q = 0$; but $\overline{C}_q \subset C_q + L_q$ is always true.

(6.3.4) L_q is also, for $q \neq p$, equal to the frontier $\text{Fr}(C_q)$.

We show that every point of L_q is a limit point of $S - \overline{C}_q$. Since $S - \overline{C}_q$ is invariant, we need this for one single point r of L_q .

We know that C_q is an open and closed set with respect to $S - L_q$; its complement in $S - L_q$ is exactly $(S - L_q) - C_q = S - (L_q + C_q) = S - \bar{C}_q$; such a complement is also open and closed in $S - L_q$. Therefore $S - \bar{C}_q$ has no limit points in C_q . It is not empty since, because of property V, $S - L_q$ is not connected whereas C_q is connected.

In S itself $S - \bar{C}_q$ is open; it could not be closed because it is neither empty nor equal to S . There must be a boundary point r , and this point is necessarily on L_q .

We shall now have to derive a series of relations between different circles and circumferences; it will be convenient to write $L_1, L_2, L_i, C_1, C_2, C_i$ instead of L_q, C_q , and so on; it is always understood that C_i is the circle determined by L_i .

(6.4) *The product L_1C_2 is either empty or L_1 .*

For L_1C_2 , as a product of invariant sets, is invariant; 0 and L_1 are the only invariant subsets of L_1 .

(6.4.1) *$L_x \subset C_y$ and $x \in C_y$ are equivalent.*

A non-trivial statement is the following:

(6.5) *$L_1C_2 = 0$ implies $C_2 \subset C_1$.*

We shall derive this by showing that the product C_1C_2 is equal to C_2 . If C_2 is empty, then $C_2 \subset C_1$ is trivially true. If not, we shall see that C_1C_2 is a non-vanishing open and relatively closed subset of C_2 ; $C_1C_2 = C_2$ follows because C_2 , as a circle, is connected.

To this purpose we determine the relative boundary of C_1C_2 in C_2 , that is, the set of all limit points of C_1C_2 which are in C_2 but not in C_1C_2 ; in other terms, the product $C_2\text{Bd}(C_1C_2)$. Here and later we shall often use the following formulas:

(6.5.1) $\text{Bd}(A+B) \subset \text{Bd}(A) + \text{Bd}(B)$; $\text{Bd}(AB) \subset \text{Bd}(A) + \text{Bd}(B)$.

Now we have $\text{Bd}(C_1C_2) \subset \text{Bd}(C_1) + \text{Bd}(C_2) \subset L_1 + L_2$; consequently,

$$C_2\text{Bd}(C_1C_2) \subset C_2L_1 + C_2L_2.$$

The set C_2L_2 is always empty, $C_2 \subset S - L_2$; C_2L_1 is empty by assumption. Hence $C_2\text{Bd}(C_1C_2) = 0$. Since C_1C_2 , absolutely open, as a product of open sets in S , is a fortiori relatively open in C_2 , it is either empty or equal to C_2 . How could C_1C_2 be empty? Only if one of the factors is empty, for otherwise both will contain the point p . The case that C_2 is empty has been disposed of; if C_1 were empty, $L_1 = \{p\}$ would imply $L_1C_2 = L_1 \neq 0$, contrary to our assumption.

Property VI has not been used yet.

(6.6) $L_1C_2 = L_1$ implies $C_1 \subset C_2$.

Considering (6.5) we see that it suffices to prove $L_2C_1 = 0$. The proof is indirect and based on property VI.

Suppose that $L_2C_1 \neq 0$; then it is equal to L_2 and $L_2 \subset C_1$.

Now consider the (open) set $C_1 + C_2$ and in particular, its boundary $\text{Bd}(C_1 + C_2)$. The relation

$$\text{Bd}(C_1 + C_2) \subset \text{Bd}(C_1) + \text{Bd}(C_2) \subset L_1 + L_2$$

implies together with

$$L_1 \subset C_2, \quad L_2 \subset C_1, \quad L_1 + L_2 \subset C_1 + C_2$$

the fact that the open set $C_1 + C_2$ contains its boundary. Hence it is equal to S or to 0. Since L_1 is in C_2 , C_2 , and a fortiori $C_1 + C_2$, are not empty, and in this way we have derived from the assumption $L_2C_1 \neq 0$ that the space S is a sum of a finite number of circles $S = C_1 + C_2$. That is excluded by property VI; hence $L_2C_1 \neq 0$ is wrong, $L_2C_1 = 0$ is true, and that implies $C_1 \subset C_2$, as we know from the preceding theorem.

As an immediate formal consequence of (6.4), (6.5), and (6.6) we obtain the next theorem:

(6.7) *If C_1 and C_2 are two circles (as always with center p), then at least one of the inclusions $C_1 \subset C_2$, $C_2 \subset C_1$ is true.*

The next theorem states the equivalence of several other inclusion relations, which we have to use later on:

(6.8) *The following properties are equivalent:*

- (a) $L_x \subset C_y$ (we know that this is equivalent to $x \in C_y$).
- (b) $\overline{C_x} \subset C_y$, and C_y is not empty.
- (c) $L_y \subset S - \overline{C_x}$, and if $x = p$ then $y \neq x$.
- (d) $C_y \not\subset C_x$.

We show that every one of these relations implies the succeeding one and that the last implies the first.

(a) *implies* (b). $L_x \subset C_y$ shows that C_y is not empty. From (6.6) we get $C_x \subset C_y$; consequently, $\overline{C_x} = C_x + \text{Bd}(C_x) \subset C_x + L_x \subset C_y + C_y = C_y$.

(b) *implies* (c). C_y is not empty; hence if $x = p$, y is not equal to x , for C_q is empty. In both cases the set $L_y \overline{C_x}$ is invariant, and hence either 0 or L_y . If it is L_y , then $L_y \subset \overline{C_x}$, $\overline{C_x} \subset C_y$ would yield the contradiction $L_y \subset C_y$. Hence $L_y \overline{C_x} = 0$ or $L_y \subset S - \overline{C_x}$.

(c) *implies* (d). In view of (6.7) let us show that $C_y \subset C_x$ is impossible. Indeed, if $x = p$, C_x is empty and $C_y \subset C_x$ would make C_y empty, whereas y is

not x . If $x \neq p$ and $y \neq p$, then $L_y \subset S - \bar{C}_x$, $L_y \subset \bar{C}_y \subset \bar{C}_x$ constitutes a contradiction. If $x \neq p$ and if $y = p$, $L_y \subset C_x$ would contradict the assumption $L_y \subset S - \bar{C}_x$.

(d) *implies* (a). From $C_y \not\subset C_x$ we infer that $C_x \subset C_y$, but not $C_x = C_y$, also that C_y is not empty, $\bar{C}_y \supset L_y$. Consequently,

$$\bar{C}_x \subset \bar{C}_y, \quad L_x \subset \bar{C}_y = C_y + L_y.$$

Hence we get for L_x

$$L_x = L_x C_y + L_x L_y.$$

The set $L_x L_y$ must be empty; for in the opposite case $L_x = L_y$, $C_x = C_y$, $C_y \subset C_x$ would ensue. It follows that $L_x = L_x C_y$, which is (a).

Abstract absolute values, symbols of the form $|x|$, where x is a point in S , and the number 0 are now introduced by the following definition:

(6.9) $|x| < |y|$ or $|y| > |x|$ shall mean $L_x \subset C_y$; $|x| = |y|$ shall mean $L_x = L_y$; $|x| = 0$ shall mean $x = p$; $|x| > 0$ shall mean $x \neq p$; $|x| \geq |y|$ shall mean $|x| > |y|$ or $|x| = |y|$.

(6.10.1) For any two points x, y exactly one of the relations $|x| < |y|$, $|x| = |y|$, $|x| > |y|$ is true.

Suppose that neither $|x| < |y|$ nor $|x| > |y|$ is true; in other terms, neither $L_x \subset C_y$ nor $L_y \subset C_x$ is true. On account of (6.4) we have then $L_x C_y = L_y C_x = 0$; from (6.5) we conclude $C_x \subset C_y$ and $C_y \subset C_x$; hence $C_x = C_y$, $L_x = L_y$, or $|x| = |y|$, which was to be shown.

(6.10.2) $|x| < |y|$, $|y| < |z|$ imply $|x| < |z|$.

We know $L_x \subset C_y$ and (cf. (6.8)) $\bar{C}_y \subset C_x$; we have a fortiori $C_y \subset C_x$; hence $L_x \subset C_x$ or $|x| < |x|$ by definition.

(6.11) $\lim x_i = p$ is true if and only if for every $|e| > 0$ an index i^* can be found such that for $i > i^*$, $|x_i| < |e|$.

For the set of all points x with $|x| < |e|$ is the circle C_e , which is, because of the relation $|e| > 0$, a neighborhood of p , and must contain almost all points of any sequence which converges towards p .

7. Geometrical version of the lemma. The following statement is of use:

(7.1) If $|x| < |y|$, then a z exists which satisfies $|x| < |z| < |y|$.

The relation $|x| < |y|$ implies, as we know, $\bar{C}_x \subset C_y$, and C_y is not empty. We maintain that $C_y - \bar{C}_x$ is not empty; for otherwise the open set C_y , neither empty nor S , would be equal to the closed set \bar{C}_x , which is impossible.

It is also impossible that $C_y - \bar{C}_x$ is equal to the one-point set $\{p\}$, for $\{p\}$ is closed and a difference "open minus closed" is open. Since $x \neq y$, the set $\{p\}$ is not equal to S . Hence we see that $C_y - \bar{C}_x$ is not only not empty but contains a point z which is not p . Any such z will do in (7.1) because $z \in C_y$ gives $L_z \subset C_y$ and $|z| < |y|$. On the other hand, z is in $S - \bar{C}_x$, hence $L_z \subset S - \bar{C}_x$; and if $x = p$, then $z \neq x$, for we took $z \neq p$; (6.8c) reveals this as an equivalent of $|x| < |z|$.

(7.1.1) *For every x there exists a y with $|y| > |x|$, if S has more than one point.*

If $|x| = 0$, take $y \neq p$; if $|x| > |p|$, take any point from $S - \bar{C}_x$, which is not empty since $S - L_x$ is not connected.

(7.2) *If $\lim x_i = x$, $\lim y_i = y$, $|x| < |y|$, then there exists an index i^* such that for $i > i^*$, $|x_i| < |y_i|$.*

Choose a z exactly as before; then $|x| < |z|$ yields $x \in C_z$, and $|z| < |y|$ implies $y \in S - \bar{C}_z$. The sets C_z and $S - \bar{C}_z$ are open; consequently, there exists an index i^* such that for $i > i^*$

$$x_i \in C_z, \quad y_i \in S - \bar{C}_z.$$

The first formula is equivalent to $|x_i| < |z|$, the second to $|z| < |y|$ since $z \neq p$. The transitive law (6.10.2) furnishes $|x_i| < |y_i|$, which was to be proved.

We may state the following corollary:

(7.2.1) *If the sequences x_i , y_i are convergent, $|x_i| = |y_i|$ for all i implies $|\lim x_i| = |\lim y_i|$.*

[7.3] *If F is a (single-valued) mapping of S in itself, then $S = S_1 + S_2 + S_3$, where S_1 , S_2 , S_3 in this order are defined by the relations $|F(x)| < |x|$, $|F(x)| = |x|$, $|F(x)| > |x|$.*

The geometric version of Schwarz' lemma is a statement about the S_i of a transformation F in N with $F(p) = p$. We derive first, with the aid of (7.2), a simple statement for continuous transformations.

(7.4) *If F in [7.3] is continuous, then S_1 and S_3 are open sets.*

We prove that S_1 is open; the proof for S_3 is virtually the same.

For a point x in S_1 we have, by definition, $|F(x)| < |x|$. Let $\lim x_i = x$; then we have to show that for almost all indices i , $|F(x_i)| < |x_i|$. This follows from (7.2) if we define $y = F(x)$, $y_i = F(x_i)$, and use the relation (continuity) $\lim F(x_i) = F(x)$.

Again we note without proof that S_2 is closed.

(7.5) *If F is a continuous mapping of S in itself and if neither S_1 nor S_3 is empty, then there exist at least two points p, q with $p \neq q$ in S_2 .*

This is a well known theorem about continuous functions coupled with the fact that S has no cut points. If S_2 were empty, $S = S_1 + S_3$ would be a non-trivial decomposition of S into two disjoint open sets, which does not exist in a connected space. If $S_2 = \{p\}$, then $S_1 + S_3$ would be a non-trivial decomposition of $S - \{p\}$ into open sets, and p would be a cut point of S .

(7.6) *Let F be in N , $F(p) = p$, such that S_2 contains p . If now S_2 contains another point q , $q \neq p$, $|F(q)| = |q|$, then F is a rotation.*

For a rotation $|F(x)| = |x|$ is identically true and S_1 and S_3 are both empty.

Proof. Since $F(q)$ and q are in the same circumference, there exists a rotation R such that $R(F(q)) = q$. What do we know about the transformation RF ? The relations $RF(p) = R(p) = p$, $RF(q) = q$ show that RF , which is in N , has two different fixed points. The corollary (4.4) tells us that RF is a rotation R_1 . From $RF = R_1$ we get $F = R^{-1}R_1$, which is a rotation since it is the product of two rotations.

(7.7) *If F is in N and $F(p) = p$, then one of the sets S_1 and S_3 is empty.*

Indeed, if none were empty there would exist two different points in S_2 (cf. statement (7.5)), and F would have to be a rotation; S_1 and S_3 would be empty.

(7.8) *If F is in N , $F(p) = p$, then S_3 is empty. In other words, $|F(x)| \leq |x|$ for all x .*

This is the geometrical version of the Schwarz lemma.

Proof. If S_3 is not empty, then S_1 is empty on account of (7.7). The set S_2 is not empty, for it contains the point p . It does not contain any others, for in that case F would be a rotation and S_3 would be empty (as well as S_1). Therefore the inequality $|F(x)| > |x|$ would hold whenever $|x| \neq 0$. But this contradicts the topological alternative (cf. (4.1)) that F is either a rotation or nilpotent. Indeed, F is not a rotation, and $|F(x)| > |x|$ for all $|x| > 0$ is incompatible with nilpotency. For $|x| > 0$ implies (by mathematical induction) $|F^n(x)| \neq 0$ and $|F^n(x)| \geq |x|$; and this would show that the Cauchy condition (6.11) for $\lim F^n(x) = p$ cannot be fulfilled with $|e| = |x|$. If our theorem were wrong, we should have a contradiction; therefore, assertion (7.8) is true.

Combining (7.6) and (7.8), we formulate the final geometrical theorem, (7.9), which corresponds to the classical Schwarz lemma together with its standard corollary.

(7.9) *If F is in N , $F(p) = p$, then for all x in S we have $|F(x)| \leq |x|$. If equality holds for one point distinct from p , then it holds throughout. In the latter case F has an inverse which is an element of N .*

The classical lemma would be a consequence provided we know that the abstract relation $|x| < |y|$ is equivalent to the analytically defined inequality $|x| < |y|$. In §8 we shall prove a theorem to the effect that the analyticity of linear homogeneous functions together with simple topological properties of the euclidean circles make the abstract and analytical order relations equivalent.

PART III

8. Characterization of circumferences. Our definitions of absolute value relations are such that if S is the unit circle in the plane of the complex numbers, p the origin, L_q the circular curve through q with center p , and C_q the interior of the corresponding circular area, then $|x| < |y|$ is equivalent to saying that the classical absolute value of x is less than the classical absolute value of y .

But we wish to know if the euclidean circumferences are circumferences in the sense of our definition. Of course it is well known that in the classical case an abstract rotation is an ordinary rotation; but this is usually shown as an application of the Schwarz lemma, or at least derived in an analogous fashion.

Let us therefore denote a euclidean circumference with K and the corresponding circle with E ; and let us discuss the case where K contains a point z but not the point p .

If we use the analyticity of linear homogeneous transformations, we see immediately that, N being the set of all analytical mappings of S in itself, K is rotatory; that is, that there exists a topological mapping in N which carries p into itself and a preassigned x on K into an arbitrary y on K . Applying some elementary topology of the euclidean plane, we can make the following assertion:

- (8.1) (a) $K \subset S$ is not empty; it contains a point z but not the point p .
- (b) $S - K$ is not connected; the component of $S - K$ which contains p is E .
- (c) K is the boundary $\text{Bd}(E)$ of E .
- (d) K is rotatory.
- (e) $\bar{E} = E + K$ is compact.

We maintain that from these statements and the properties I-VI of N and I-III of S it follows that K is a circumference. (The case $K = \{p\}$ is trivial, since $L_p = p$.)

Let us forget the euclidean origin of (8.1) and make the following definition:

[8.2] " (K, E, z) is circular" shall mean that the sets $K \subset S$, $E \subset S$, and the point z satisfy the relations (8.1).

The "justification theorem" in question is now simply the following:

(8.3) Let N and S be as in Part II. If (K, E, z) is circular, then K is a circumference and E a circle; in short $K = L_z$, $E = C_z$.

Due to the definition (in (8.1a)) of E and [5.2] of C_z it is sufficient to show $K = L_z$.

The proof is arranged backwards:

(8.4) If (K, E, z) is circular and if no point of the circumference L_z is in E , then $K = L_z$.

Consider the set $C_z E$; this set, the product of two open sets, is open. (The set E is open since K , being a boundary, is closed.) The set $C_z E$ is not empty because p is in C_z and in E .

We study, as we always did in questions of this type, the relative boundary of $C_z E$, this time with respect to both C_z and E .

Note that K is a subset of L_z , for it contains z and is rotatory. We get $\text{Bd}(C_z E) \subset \text{Bd}(C_z) + \text{Bd}(E) \subset L_z + K \subset L_z$. Of course we cannot conclude directly that equality holds, for we do not know yet that K is invariant. But at least we can say that the relative boundaries $C_z \text{Bd}(C_z E)$ and $E \text{Bd}(C_z E)$ are empty. That $C_z L_z = 0$ follows from the definition of C_z ; whereas $L_z E = 0$ is an assumption of our theorem. Hence the relative boundaries of $C_z E$ with respect to the (connected) sets C_z and E are empty as subsets of $C_z L_z$ and $E L_z$, respectively. Since $C_z E$ is not empty, we obtain $C_z E = C_z$, $C_z E = E$; hence $C_z = E$. Taking boundaries on both sides, we have $L_z = K$, which was to be proved.

(8.5) If (K, E, z) is circular and if S is not compact, then L_z has no point in common with E .

The proof is indirect: If q is in $L_z E$, then $L_q = L_z$, and to every point x in $L_q = L_z$ there will exist a rotation $R^{(x)}$ such that $R^{(x)}(q) = x$. The open set E is transformed into open sets $R^{(x)}(E)$, and $q \in E$ implies $R^{(x)}(q) = x \in R^{(x)}(E)$. In other terms,

$$L_z \subset \sum_{x \in L_z} R^{(x)}(E).$$

Now we have to use, for the first time, the metrizable of the space S . Since

L_z is a compact subset (cf. (3.5)) of a metrizable space, the Heine-Pincherle-Borel-Lebesgue† theorem is valid, and already a finite number of sets $R^{(z)}(E)$, say $R_1(E), \dots, R_n(E)$ covers L_z ; that is,

$$L_z \subset \sum_1^n R_i(E).$$

We set $S' = \sum_1^n R_i(E)$ and propose to show that $S' = S$. This is again done with the standard device based on the connectedness of S .

The set S' is open as a sum of open sets; it is not empty because it contains L_z . What is its boundary? We obtain

$$\text{Bd}(S') = \text{Bd}\left(\sum_1^n R_i(E)\right) \subset \sum_1^n \text{Bd}(R_i(E)) = \sum_1^n R_i(\text{Bd}(E)) \subset \sum_1^n R_i(L_z) \subset L_z.$$

(We have applied (6.6.1), (6.2.4), $K = \text{Bd}(E)$, $K \subset L_z$, and $R_i(L_z) \subset L_z$.) Isolating the first and the last terms, we have $\text{Bd}(S') \subset L_z$; and since $L_z \subset S'$ we see that the open, non-empty set S' contains its boundary; S is connected, hence (cf. (6.3.1)) $S' = S$.

From $S = \sum_1^n R_i(E)$ we obtain a fortiori $S = \sum_1^n R_i(\bar{E})$. Since (K, E, z) is circular, \bar{E} and its topological images $R_i(\bar{E})$ are compact; the sum of a finite number of compact sets is compact; hence S is compact, which contradicts the assumption of the theorem. Hence we have seen, indirectly, that if S is compact, $L_z E = 0$, which was to be shown.

Finally, we remove, in (8.6), the last condition.

(8.6) S is not compact.

For if it were, it would have to be bicomact, being metrizable. Consider the covering which is defined by assigning to p the open set C_p and to every other point x the circle with center x determined by p . If S were bicomact, a finite number of these circles would have the sum S , which is excluded by property VI.

With (8.6) the proof of the justification theorem (8.3) is completed.

9. Separability and local compactness of S . If we use the foregoing theory for variable centers p , we see that every point is contained in arbitrarily small neighborhoods with compact, metrizable and hence separable boundaries. From a theorem of F. B. Jones‡ we could infer the next theorem:

† A space S is called bicomact or the Heine-Pincherle-Borel-Lebesgue theorem holds in S if from every covering of S by open sets a finite set of elements (open sets) can be extracted which has S as its sum.

‡ F. B. Jones, *A theorem concerning locally peripherally separable spaces*, Bulletin of the American Mathematical Society, vol. 41 (1936), p. 437.

(9.1) *The space S is (perfectly) separable.*

This result will also appear as a corollary of the theorem (9.8). Independently from (9.1) we are going to show that the closed circles \bar{C}_i are compact, and that the space S is representable as the sum of a countable number of circles.

(9.2) *Let x_i be a sequence of points such that a point x , a subsequence x'_i and a sequence of rotations R_i can be found with $\lim R_i(x'_i) = x$. Then there exists also a limit point for the sequence x_i .*

We select corresponding subsequences R'_i, x''_i such that $\lim R'_i = R$ exists; we know then that $\lim R'^{-1}_i = R^{-1}$, and from $\lim R'_i(x''_i) = x$ it follows that $x''_i = R'_i(R'^{-1}_i(x''_i))$ is convergent (with the limit $R^{-1}(x)$).

(9.3) *Suppose that the sequence x_i is such that sequences x'_i, R_i , as described in (9.2), do not exist. Then for every point y in S there exists a neighborhood U_y and an index i^* such that for $i > i^*$, U_y is completely in C_{x_i} or completely in the exterior of C_{x_i} .*

In other terms, $i > i^*$ implies that either $U_y \subset C_i$ or $U_y \subset S - \bar{C}_{x_i}$.

As before, we shall write C_i for C_{x_i} , L_i for L_{x_i} .

We first choose a neighborhood V_y and an index i^* such that for all indices $i > i^*$ $L_i V_y = 0$, and in addition for $i > i^*$, $C_i \neq 0$. Such a V_y exists; for L_i consists exactly of the points $R(x_i)$, where R is arbitrary. With the first countability axiom of Hausdorff (a trivial consequence of the metrizability of S) a sequence $R_i(x'_i)$ with limit y could be constructed. If $C'_i = 0$ for a subsequence C'_i , then $x'_i = p$, $R_i = I$ yields $\lim R_i(x'_i) = p$.

Since S is locally connected, V_y contains a connected neighborhood U_y of y . Now consider the formula

$$U_y = U_y \bar{C}_i + U_y(S - \bar{C}_i);$$

if $i > i^*$, then $C_i \neq 0$ and $\bar{C}_i = C_i + L_i$; hence

$$U_y = U_y C_i + U_y L_i + U_y(S - \bar{C}_i).$$

For $i > i^*$, $U_y L_i$ is 0; the resulting equation

$$U_y = U_y C_i + U_y(S - \bar{C}_i)$$

is a decomposition of U_y into two disjoint open sets. One of these must be empty, since U_y is connected and not empty; but that means that either $U_y \subset C_i$ or $U_y \subset S - \bar{C}_i$, which was to be shown.

(9.4) *Let the sequence x_i be such that to every y in S there belongs a neighborhood U_y and an index i^* such that for $i > i^*$ either $U_y \subset C_i$ or $U_y \subset S - \bar{C}_i$ is true. Then the set $S' = \sum C_i$ (which is trivially open) is closed.*

In order to prove this we show that if y_i is a convergent sequence from S' , its limit y is also an element of S' .

Without loss of generality we may assume $y_i \in C_i$; this corresponds to the deletion of some C 's and introduction of a new index, which does not affect the validity of our theorem.

Now let i^* be such that for $i > i^*$ (a) $y_i \in U_y$ and (b) either $U_y \subset C_i$ or $U_y \subset S - \bar{C}_i$.

(a) may be satisfied since $\lim y_i = y$; (b) has been explicitly assumed. Since for $i > i^*$, $y_i \in U_y$, $y_i \in C_i$, we see that $y_i \in U_y C_i$; that decides the alternative (b) in favor of $U_y \subset C_i$; but $U_y \subset C_i$ (any special case such as $i = i^* + 1$ will do) implies $y \in C_i$ and a fortiori $y \in \sum C_i = S'$.

(9.5) *If a sequence x_i has no limit point, then $S = \sum C_{x_i}$ ($= \sum C_i$).*

Since $x_i = p$ can be true only a finite number of times, almost all C_i are not empty; a fortiori the open set $S' = \sum C_i$ is not empty.

(9.2), (9.3), (9.4) together guarantee that S' is closed; the connectedness argument yields $S' = S$.

(9.6) *The circles C_q are limited, and their closures \bar{C}_q compact.*

The proof is indirect. Let x_i be a sequence from C_q . If it had no limit point, we would have $S = \sum C_{x_i}$. On the other hand, $x_i \in C_q$ implies $C_{x_i} \subset C_q$, (cf. (6.6)) since $C_q \neq 0$. That would lead to the contradiction $S \subset C_q$, since q is not in C_q . Hence every sequence x_i from C_q must have a limit point, which was to be proved.

We could express and slightly generalize this in the following familiar form:

(9.6.1) *If all points x of a set satisfy $|x| \leq |q|$, then every infinite sequence x_i has a limit point.*

As a consequence we have the statement:

(9.7) *S is locally compact.*

For if x is an arbitrary point, there exists a point y such that $|x| < |y|$, $x \in C_y$. Hence every point is contained in a limited open set.

(9.8) *S is semicompact; that is, it is the sum of a sequence of compact sets \bar{C}_i . (They will be closed circles.)*

If S were compact (which is excluded by (8.6)), then it would be trivially semicompact. If it is not, then there exists a sequence x_i without limit points. In that case we have (cf. (9.5)) $\sum C_{x_i} = S$ and a fortiori $\sum \bar{C}_i = S$; and the \bar{C}_i are now known to be compact.

From the theorem (9.8) (all theorems in this section are proved without recourse to (9.1)) we get (9.1) as a trivial consequence, using metrizability.

Conclusion. It would be possible to obtain valuable new properties of S and N by adjunction of new postulates. We could demand that the circumferences be connected; this would permit us to conclude that for every pair x, y a center p and a rotation R exist such that $R_p(x) = y$. If we postulate that a transformation with two fixed points is the identity, L_q would be homeomorphic to a connected compact continuous group. These groups are rather well known, and together with the fact that the abstract absolute values can be interpreted as real numbers, this additional axiom would heavily restrict the structure of the space S . If, finally, S is supposed to be homeomorphic to the euclidean plane, the application of a theorem of Hilbert would show that the invertible transformations in N induce an absolute, that is, either euclidean or hyperbolic, geometry in S . The decision as to whether these axioms together with a maximality axiom are categoric will largely depend, we believe, on the better understanding and proper generalization of the Schwarz lemma.

UNIVERSITY OF CALIFORNIA AT LOS ANGELES,
LOS ANGELES, CALIF.

STEINITZ FIELD TOWERS FOR MODULAR FIELDS*

BY
SAUNDERS MAC LANE

1. **Introduction.** The systematic study of the most general modular fields of characteristic p appears in its classical form in the famous Steinitz monograph [5]. Very little further analysis of such fields has been undertaken, except that in 1934 Hasse and Schmidt [2] showed that the structure of complete fields with valuations can be discussed in terms of a suitable transfinite but "separable" generation for an arbitrary modular field K . The theorem that such a "separable" generation (with the specific properties quoted below, in §9) must exist for every field K they stated but did not prove. We propose to show that their theorem, as stated, cannot be true. First, certain special cases or modifications of this theorem can be established, as in §§4 and 5, but badly imperfect fields and fields obtained by the adjunction of a denumerable infinitude of algebraically independent elements can be suitably constructed (§§7 and 8) as counter-examples to the general theorem. The most elaborate of our counter-examples, given in §8, seems almost pathological, but actually initiates many problems on the structure of such modular fields, such as the generalization of the lemmas used to analyze such an example or the formulation of other canonical generations for arbitrary fields.

What "separable" generations of a field K are considered? If K can be obtained from a prime field P by the successive adjunction of elements, each one of which is transcendental or separable algebraic over the field previously obtained, then K has a "separating transcendence basis" over the subfield P . When there is no such separating basis, it may still be possible to represent the whole field K as the union of the fields of a tower

$$(1) \quad M_0 \subset M_1 \subset M_2 \subset \cdots \subset K,$$

in which each individual field M_i does have a separating transcendence basis. Such towers of "residue-class fields" M_i appear in the Hasse-Schmidt analysis of a topologically complete field \mathfrak{K} with a discrete valuation. The "residue-class field" K of such a field \mathfrak{K} is obtained just as the Galois field of p elements is obtained by reducing the integers (or the p -adic integers†) modulo p . To construct a complete field \mathfrak{K} with given residue-class field K one seeks to obtain \mathfrak{K} by successive extensions

* Presented to the Society, November 26, 1938; received by the editors October 14, 1938.

† For a discussion of valuations, see, for instance, Albert [1, chaps. 11 and 12].

$$(2) \quad \mathfrak{M}_0 \subset \mathfrak{M}_1 \subset \mathfrak{M}_2 \subset \cdots \subset \mathfrak{K}$$

parallel to the tower (1) under the residue-class field. This parallel construction requires the Hensel-Rychlik irreducibility theorem, which states that a separable polynomial $g(t)$ with a root x in the residue-class field K corresponds to a polynomial over the complete field \mathfrak{K} with a root ξ in \mathfrak{K} (and in the residue class x).^{*} This theorem will construct $\mathfrak{M}_1/\mathfrak{M}_0$ provided M_1 is separable and algebraic over M_0 in (1); hence the desirability of a separable tower (1).

For a perfect field K , Schmidt obtained such a "Steinitz" separating tower. For instance, if $K = P(t, t^{p^{-1}}, t^{p^{-2}}, \dots)$ is obtained from a perfect subfield P by the adjunction of all p th roots of a single indeterminate t , and if S_e is the subfield $P(t^{p^{-e}})$, then we have a "tower"

$$S_0 \subset S_1 \subset S_2 \subset S_3 \subset \cdots, \quad K = \sum_{e=0}^{\infty} S_e,$$

where the summation sign used here denotes the "union" or "composite" of the fields S_e indicated. This tower then has additional properties:

- (i) Each S_e is separable over the transcendence basis $t^{p^{-e}}$.
- (ii) $S_{e-1} = S_e^p$, where S_e^p denotes the field of all p th powers of elements of S_e .

For imperfect fields Schmidt has formulated a generalized Steinitz tower, constructed over a suitable base field L , and having properties similar to (i) and (ii). Our counter-examples[†] concern this tower, and we show in §§4 and 5 that a modified such tower is possible if K has a finite transcendence basis over L . Our chief tool is Lemma I in §2, which makes it possible to exchange certain elements for other elements in a given transcendence basis, without any loss of separability. This lemma resembles the so-called Steinitz exchange theorem.

The subfields L over which the towers for the "relatively perfect" field K are to be constructed are obtained in §3 by a simple application of Teichmüller's notion of the p -basis of a modular field [6, §3]. The last paragraphs contain a precise statement of the relation of our counter-examples to the theorem of Schmidt.

2. The exchange lemma. We consider exclusively fields of characteristic a fixed prime p . If $L \subset K$ are such fields, an element α of K is said to be

^{*} Cf. Hasse-Schmidt [2, p. 31], or, for the p -adic number case, Albert [1, Lemma, p. 296].

[†] The results of Hasse-Schmidt in [2] on complete fields with valuations are not called into question, since Witt and Teichmüller have subsequently established them by other methods. See [7], [8], or [4].

separable over L if α satisfies over L an irreducible polynomial equation without multiple roots. The field K is *separable algebraic* over L if every element of K is separable and algebraic over L . If K is not algebraic over L , a *transcendence basis* for K over L is a subset T of K such that K is algebraic over $L(T)$ but not algebraic over $L(T')$ for any proper subset T' of T . Here $L(T)$ denotes the field obtained from L by adjoining all elements of the set T .

We shall be concerned often with a "separating" basis T . A subset T of K is a *separating transcendence basis* (s.t.b.) for K over L if and only if T is a transcendence basis for K over L such that K is separable algebraic over $L(T)$.

THEOREM 2.1. *A field K has a separating transcendence basis over a subfield L if and only if the elements of K can be well ordered in such a way that every element b of K is either transcendental or separable algebraic over the field K_b obtained by adjoining to L all elements prior to b in the well ordering of K .*

For a given s.t.b. T the required well ordering can be constructed by listing first the elements of T in any order and then the remaining elements of K in any order. Conversely, given the well ordering, the corresponding s.t.b. T is simply the set of those elements b which are, respectively, transcendental over the corresponding fields K_b . A field K with such a well ordering has been called by Schmidt [2] a field "separable" over L . Hence K is "separable" over L if and only if it has a s.t.b.

Inseparable equations involve the variables only as p th powers. If a polynomial in the variable y (the coefficients may involve other variables) can be written as $f(y) = \sum a_i y^{ip^e}$ with at least one $a_i \neq 0$ we say that f has *exponent* p^e in y . We recall that an element α inseparable over a field L satisfies an irreducible equation $f(y) = 0$ over L in which y has exponent $p^e > 1$; furthermore p^e is the smallest exponent such that α^{p^e} is separable over L . We call p^e the *exponent* of α over L . (In Steinitz' work e itself was known as the exponent.)

The following "exchange" lemma is used repeatedly:

LEMMA I. *If in a field K the elements of a subset $T \subset K$ are algebraically independent* over a perfect subfield P of K , and if the element y of K is separable over the field $P(T)$, while $y^{1/p}$ is not separable over $P(T)$, then there is an element x in the set T such that y is not separable over $\dagger P(T - \{x\}, x^p)$. Any such element x is separable over $P(T - \{x\}, y)$, but not over the field $P(T - \{x\}, y^p)$.*

In effect, the lemma says that the fields $P(T - \{x\}, x)$ and $P(T - \{x\}, y)$ each consist of elements separable over the other field—an exchange of x for y .

* A set of elements is algebraically independent over a field if the elements satisfy no non-trivial polynomial equations with coefficients in the field.

† Here $T - \{x\}$ denotes the set T with the element x deleted.

Proof. The algebraic equation for y over $P(T)$ can be written in the form $g(y, T) = 0$, where g has coefficients in P , is of exponent 1 in y , and is irreducible as a polynomial over P in the variables y, T . If g had exponent p or greater in each variable of T , we could take the p th root of each term in $g(y, T)$ to get a separable equation for $y^{1/p}$ over $P(T)$, counter to hypothesis. Therefore, at least one quantity x of T appears in g with exponent 1. If $T' = T - \{x\}$ be the set of the remaining elements in T , the equation $g = 0$, in the form $g(y, x, T') = 0$, shows that x is algebraic over the field $P(y, T')$. The elements y, T' of the set generating this field are therefore algebraically independent.

By construction, $g(y, x, T')$ is irreducible as a polynomial in the variable x over the ring $P[y, T']$. The Gauss lemma shows that $g(y, x, T')$ is also irreducible as a polynomial in x over $P(y, T')$. Since the polynomial has exponent 1 in x , the root x is separable over the field $P(y, T')$, as asserted. Furthermore, x cannot be separable over the smaller field $P(y^p, T')$, for in that event y would be separable over $P(x, T')$ which in turn would be separable over $P(y^p, T')$, although y manifestly satisfies an inseparable irreducible equation of degree p over $P(y^p, T')$.

The element x so exchanged with y was chosen as any element of T of exponent 1 in the equation $g(y, T) = 0$. The assertion of the lemma that it may be chosen as any x such that y is not separable over $P(T - \{x\}, x^p)$ is a result of the following lemma:

LEMMA II. *If the elements of $T \subset K$ are algebraically independent over a perfect subfield P of K , and if an element y in K satisfies a separable polynomial equation $f(y) = 0$ with coefficients in $P[T]$ and irreducible over $P[T]$, then an element x of T appears in this equation with exponent 1 if and only if y is inseparable over $P(T - \{x\}, x^p)$.*

If x appears in f only with exponent $p^e > 1$, then $f(y) = 0$ is manifestly an irreducible separable equation for y over $P(T', x^p)$, where $T' = T - \{x\}$. This establishes one half of the lemma. Conversely, suppose that x appears with exponent 1 in $f(y)$. Then $f(y) = g(y, x, T')$ has exponent 1 in x and in y and is irreducible in the ring $P[y, x, T']$, where y, x , and T' are regarded as independent variables. Consider the polynomial

$$g^{(p)}(y^p, x^p, T'^p) = [g(y, x, T')]^p$$

where $g^{(p)}$ denotes the function obtained from g by replacing each coefficient by its p th power. Then $g^{(p)}(y^p, x^p, T'^p)$, which is the p th power of an irreducible polynomial g in $P[y, x, T']$, can be in no way reducible in the smaller ring $P[y, x^p, T']$, which does not contain this irreducible factor $g(y, x, T')$. In other words, $g^{(p)}(y^p, x^p, T'^p)$ is irreducible in $P[y, x^p, T']$ and hence by the

Gauss lemma is irreducible in $P(x^p, T')[y]$. This means that the element y satisfies an equation with exponent p over $P(x^p, T')$, which makes y inseparable over this field $P(x^p, T')$, as asserted.

An element α is said to be *purely inseparable* over a field L if α^{p^m} is in L for some power p^m . If p^m is chosen as the least such power, then $x^{p^m} - \alpha^{p^m} = 0$ is known to be the irreducible equation satisfied by α over L . A field K is purely inseparable over L if every element of K is purely inseparable over L . If an element α of K is both purely inseparable and separable algebraic over L , then α satisfies over L two equations, a separable equation $f(x)$ with no multiple roots and a purely inseparable equation with only one root. The greatest common divisor of these two equations is linear and has the form $x - \alpha = 0$, with a coefficient α in L ; hence the useful remark (Teichmüller [6, Theorem 12]):

LEMMA III. *An element α both separable and purely inseparable over a field L lies in that field.*

3. Relatively perfect intermediate fields. The *perfect closure* or *least perfect extension* of a field K is the field obtained by adjoining to K all roots x^{1/p^e} of elements x in K , for all integers $e \geq 0$. If K^{p^e} is taken to denote the field of all elements x^{p^e} , for x in K , then $K^{p^{-1}}$ is the field obtained from K by the adjunction of all p th roots of elements of K , while the perfect closure $K^{p^{-\infty}}$ becomes $K^{p^{-\infty}} = K(K^{p^{-1}}, K^{p^{-2}}, \dots)$.

F. K. Schmidt has called a field K *relatively perfect* over a subfield L if the perfect closure of K can be obtained by adjoining to K roots of elements in L alone; that is, if $K^{p^{-\infty}} = K(L^{p^{-\infty}})$. Here $K(L^{p^{-\infty}})$ can be considered as the composite $K \cup L^{p^{-\infty}}$ of K and $L^{p^{-\infty}}$ formed within the larger field $K^{p^{-\infty}}$. In particular, K is certainly relatively perfect over L if $K = K^p(L)$; that is, if $K^{p^{-1}} = K(L^{p^{-1}})$. For the construction of field towers we use the existence of such subfields L in the following explicit sense:*

THEOREM 3.1. *If P is a perfect subfield of K , then there exists an intermediate field L with $P \subset L \subset K$ such that $K = K^p(L)$ and such that L has a separating transcendence basis over P and is relatively algebraically closed† in K .*

To establish this theorem, we utilize the notion of p -independence due to Teichmüller [6]. A subset X of K is p -independent in K if $K^p(X')$ is a proper subfield of $K^p(X)$ whenever X' is a proper subset of X . Alternatively, X is p -independent if and only if no element x in X is contained in the field $K^p(X - \{x\})$. A subset X of K is a p -basis of K if X is p -independent in K

* F. K. Schmidt [2] states without proof a similar theorem, omitting the property, essential to our purposes, that L is relatively algebraically closed in K .

† L is relatively algebraically closed in K if and only if every element of K algebraic over L is in L .

and if, in addition, $K = K^p(X)$. It follows readily that X is p -independent in K if and only if each finite subset of X is p -independent. This means, in other words, that the degree $[K^p(x_1, \dots, x_m):K^p]$ is p^m for any m distinct elements x_1, \dots, x_m of X . The latter statement was used as a definition of p -independence by Teichmüller [6, §3], so that our definition agrees with his. We next obtained another alternative definition based on the following:

LEMMA 3.2. *If Y is a p -independent subset of K , then $K \cap K^p(Y) = K \cap K^p(Y^{p^{-\infty}})$.*

Here and subsequently $K \cap L$ denotes the intersection of the fields K and L , while $K^p(Y^{p^{-\infty}})$ designates the field $K^p(Y, Y^{p^{-1}}, \dots)$ obtained by adjoining to K^p all elements $y^{p^{-e}}$, for y in Y and e a positive integer.

Proof. We need only derive a contradiction from the assumption that some x of K not in $K^p(Y)$ is in $K^p(Y^{p^{-\infty}})$. For such an x there is an integer $e > 0$ such that x is in $K^p(Y^{p^{-e}})$, but not in the field $M_e = K^p(Y^{p^{-e+1}})$. There then is a finite subset Z of Y such that x is in the field $M_e(Z^{p^{-e}})$. Therefore x has the form $x = f(y_1^{p^{-e}}, \dots, y_n^{p^{-e}})$ where each y_i is an element of Z , where the polynomial f has coefficients in M_e , has degree less than p in each variable $y_i^{p^{-e}}$, and contains at least one variable, say $y_1^{p^{-e}}$, with an exponent 1. If g is the polynomial obtained from f by replacing each coefficient by its p^e th power, then

$$(1) \quad x^{p^e} - g(y_1, \dots, y_n) = 0$$

where g has coefficients in $M_e^{p^e} \subset K^p$, and is of degree less than p in y_1 . Hence, over the field $K^p(y_2, \dots, y_n)$, y_1 satisfies the separable equation (1) as well as the purely inseparable equation $y_1^p = a$, a in K^p . Therefore y_1 lies in $K^p(y_2, \dots, y_n)$ as in Lemma III, contrary to the assumed p -independence of the set Y .

From this lemma one obtains the following theorems:

THEOREM 3.3. CRITERION FOR INDEPENDENCE. *A subset X of K is p -independent in K if and only if no x in X is contained in the field $K^p(X_0^{p^{-\infty}})$ where $X_0 = X - \{x\}$ is the set X with x deleted.*

THEOREM 3.4. *A subset X of K is a p -basis of K if and only if X is a p -independent subset of K for which $K^{p^{-\infty}} = K(X^{p^{-\infty}})$.*

Proof. If X is a p -basis, then by definition $K = K^p(X)$, so that an application of the isomorphism $a \mapsto a^p$ yields the equation $K^p = K^{p^2}(X^p)$. By induction, we then obtain $K = K^{p^e}(X)$, or, by another isomorphism carrying each element into its p^e th root, $K^{p^{-e}} = K(X^{p^{-e}})$. This yields the conclusion that $K(X^{p^{-\infty}})$ is the perfect closure of K .

Conversely, if $K^{p^{-\infty}} = K(X^{p^{-\infty}})$, then $K^{p^{-\infty}} = K^p(X^{p^{-\infty}})$, $K \subset K^p(X^{p^{-\infty}})$, and hence by Lemma 3.2, $K \subset K^p(X)$. This is exactly the condition used to define a p -basis.

Returning to the existence of relatively perfect subfields, we prove a more explicit form of Theorem 3.1.

THEOREM 3.5. *If P is a perfect subfield of K , if X is any p -basis of K , and if L is the field of all elements of K algebraic over $P(X)$, then X is a separating transcendence basis for L over P and $K = K^p(L)$.*

When this theorem has been established, Theorem 3.1 will be an immediate consequence, for a straightforward argument by transfinite induction can be used to establish the existence of a p -basis X for any field K (Teichmüller [6]).

Proof. The fact that the set X is algebraically independent over P is known (Teichmüller [6, Theorem 15]). If L did not have X as a s.t.b., there would be an element z in L inseparable with exponent p over $P(X)$.

The element $y = z^p$ is therefore separable over $P(X)$, although $y^{1/p}$ is not so separable, as in the hypothesis of Lemma I (§2). The conclusion of that lemma produces an element x in X which is separable over $P(X - \{x\}, z^p)$ and hence over the larger field $K^p(X - \{x\})$. But x is also purely inseparable over $K^p(X - \{x\})$, and therefore x must be contained in the field $K^p(X - \{x\})$, contrary to the assumed p -independence of the set X .

Finally, since $X \subset L$ is a p -basis of K , $K = K^p(X) \subset K^p(L)$ must hold, as stated in the theorem.

4. The Steinitz field tower. Throughout this section we shall study the properties of a certain tower of fields over one of the intermediate fields L constructed in the last theorem.

HYPOTHESIS. P is a perfect subfield of K ; X is a p -basis of K ; L is the field of elements of K algebraic over $P(X)$.

For any transcendence basis T of K over L , we consider the set

$$(1) \quad S_n = \mathfrak{S}_n(K; L(T)) = [\text{all } \alpha \text{ in } K \text{ with } \alpha^{p^n} \text{ separable over } L(T)],$$

consisting of all elements of K with exponents p^n or less over $L(T)$. Steinitz [5, §14, Theorem 2] showed that S_n is a field and that K is the union of these fields S_n :

$$(2) \quad S_0 \subset S_1 \subset S_2 \subset \cdots; \quad K = S_0(S_1, S_2, \cdots).$$

We call this chain of fields a Steinitz field tower for K over L . Steinitz' results also yield (Steinitz [5, §13, Theorem 1]) the following description of this tower:

LEMMA 4.1. *Each field S_n of the tower (2) consists of those elements of K of exponent p or less over the previous field S_{n-1} .*

If $K > L$, then T is non-void. Furthermore each inclusion in the tower (2) is a proper inclusion. For were $S_n = S_{n-1}$, there would be no elements of exponent p^n and hence no elements of any larger exponent over $L(T)$. Therefore $K = S_{n-1}$ and $K^{p^{n-1}} \subset S_0$, which means that $K^{p^{n-1}}$ is separable over $L(T)$, while K^{p^n} is separable over $L(T^p)$. Because X is a p -basis of K , the definition of §3 makes $K = K^{p^n}(X) = K^{p^n}(L) = L(K^{p^n})$. This implies that K , like K^{p^n} , is separable over $L(T^p)$, and that any t in T is so separable. But $t = (t^p)^{1/p}$ is also purely inseparable over $L(T^p)$ so that Lemma III requires t to be in $L(T^p)$. This is a contradiction because the set T^p is composed of elements algebraically independent over L . We conclude that

$$(3) \quad K > L \text{ implies } S_n > S_{n-1}, \quad n = 1, 2, \dots$$

In the special case K a perfect field, the structure of the Steinitz tower has been formulated thus by Schmidt:

THEOREM OF F. K. SCHMIDT. *If K is a perfect field containing a perfect field $L = P$ relatively algebraically closed in K and if K has a transcendence basis T over L , then*

(i) *The n th field S_n of the Steinitz tower (2) has the separating transcendence basis $T^{p^{-n}}$ over P ;*

(ii) $S_n = P(S_{n+1}^p)$.

Proof. In this case, we can assume $L = P$ because L is constructed from a p -basis X , whereas a p -basis of a perfect field is automatically empty. The second conclusion of the theorem can be asserted in the stronger form $S_n = S_{n+1}^p$ because of Lemma 4.1 and because each element of S_k has a p th root in the perfect field and hence in the field S_{n+1} . Furthermore, if y is an element of S_n , then y^{p^n} satisfies a separable irreducible equation with coefficients polynomials from $P[T]$, so that the p^n th root of this equation yields for y itself a separable equation with coefficients in $P[T^{p^{-n}}]$. Therefore $T^{p^{-n}}$, patently contained in S_n , is a s.t.b. for S_n , as asserted.

Our main problem is then the investigation of the two properties (i) and (ii) given for the Steinitz tower in this theorem, in the case K not a perfect field. We consider first the question of separating transcendence bases as in property (i). Our next objective is the following theorem:

THEOREM 4.2. *If K has a finite degree of transcendence over L , then each field S_n of the Steinitz field tower (2) has a separating transcendence basis T_n over L and hence also has a separating transcendence basis $X + T_n$ over P . Each basis T_n has the same number of elements as does T .*

Proof. We construct first a transcendence basis for S_1 . Suppose that the finite transcendence basis T has exactly m elements which are p th powers in K , so that

$$(4) \quad T = U + W^p, \quad W^p = \{w_1^p, \dots, w_m^p\},$$

while no element of U is in K^p . Then $U + W$, where W is the set $\{w_1, \dots, w_m\}$, consists of elements of the field S_1 . If this set $U + W$ is not already a s.t.b. for S_1 , there is an element z in S_1 not separable over $L(U + W)$. We seek a modified basis T^* containing $y = z^p$. By hypothesis, the element z has exponent p over $L(T) = L(U, W^p)$ and also over $L(U, W)$. Since L is separable over $P(X)$, K/P has the transcendence basis $X + U + W$, and, by the transitivity of separability, z has exponent p over $P(X, U, W^p)$ and $P(X, U, W)$.

Let $f(z) = 0$ be the irreducible equation for z over the polynomial ring $P[X, U, W]$. Then f must have exponent p in z ; but no element of W can appear in f with an exponent 1, for otherwise Lemma II would imply that z^p is inseparable over $P(X, U, W^p)$, contrary to hypothesis. Suppose that all the variables of U appear with exponent at least p in f . As f is irreducible and inseparable in z , at least one of the elements of $X + W + U$ has exponent 1 in f . This must then be an element x of X . Since $f(z)$ is irreducible over $P[X, U^p, W^p]$, Lemma II implies that $y = z^p$ is inseparable over the field $P(X_0, U^p, W^p, x^p)$ where $X_0 = X - \{x\}$. Therefore, by Lemma I, x is separable over $P(X_0, U^p, W^p, z^p)$, and hence over $K^p(X_0)$. This contradicts the assumed p -independence of X .

There must then be an element u from U with exponent 1 in $f(y)$; in particular, we know that U is not void. Another application of the exchange lemma to the polynomial $f(y)$ shows that u is separable over $P(X, W^p, U_0, z^p)$, where $U_0 = U - \{u\}$. In other words, the transcendence basis

$$(5) \quad T^* = U_0 + W^p + \{z^p\}, \quad U_0 = U - \{u\},$$

for K over L has exactly $m+1$ p th powers, one more than T , and T^* is separably equivalent to T in the sense that $L(T)$ is separable over $L(T^*)$ and conversely. Consequently $\mathfrak{S}_n(K; L(T)) = \mathfrak{S}_n(K; L(T^*))$ for every n , so T and T^* yield the same Steinitz towers (2).

Repeated applications of this transition from T to T^* whenever $U + W$ is not already a s.t.b. for S_1 will, after a finite number of steps, either yield a s.t.b. for S_1 or a new transcendence basis $T_r = W_r^p$ for K/L consisting only of p th powers. In this case the remark above that $U \neq 0$ shows that W_r must be a s.t.b. T_1 for all z in S_1 .

This construction of a basis T_1 for S_1 yields by induction a similar s.t.b. for each S_n , for according to Lemma 4.1, S_n consists of elements of exponent p

or less over S_{n-1} , just as S_1 consists of elements of exponent p or less over S_0 . The theorem is thus established.

For a subsequent use in §5 we need the following lemma:

LEMMA 4.3. *If $C_n = S_n - S_{n-1}$ is, for $n > 0$, the set of all elements of K of exponent exactly p^n over $L(T)$, then, when $T \neq 0$,*

$$L(C_n^p) = L(S_n^p), \quad L(C_n) = L(S_n).$$

Proof. By (3), there exists an element x in C_n with exponent p^n over $L(T)$. If y is an arbitrary element of S_n not in C_n , then y has an exponent p^m , ($m < n$), over $L(T)$. Hence y lies in S_{n-1} and xy must be in $C_n = S_n - S_{n-1}$. Since x is in C_n , y is in $L(C_n)$; therefore $L(S_n) \subset L(C_n)$. Similarly

$$(xy)^p \in C_n^p, \quad x^p \in C_n^p, \\ y^p = (xy)^p / x^p \in L(C_n^p), \quad L(S_n^p) \subset L(C_n^p).$$

5. Modified towers of fields. When K is itself a perfect field, the Steinitz field tower (§4, (2)) has the useful property (ii) of Schmidt's theorem (§4): $S_n = P(S_{n+1}^p)$. Though we cannot assert this fact for every Steinitz field tower, we can in certain cases obtain another tower with an analogous property by omitting certain of the fields from the Steinitz tower.

THEOREM 5.1. *If, in the hypothesis of §4, the transcendence basis T for K over L is finite, then there exists a set of subfields M_k of K ,*

$$(1) \quad M_0 \subset M_1 \subset M_2 \subset M_3 \subset \cdots \subset K, \quad K = \sum_{k=0}^{\infty} M_k,$$

where \sum denotes the union of the fields M_k , such that

- (i) Each M_k has a separating transcendence basis T'_k over L ,
- (ii) $M_k \subset L(D_{k+1}^p)$ where $D_{k+1} = M_{k+1} - M_k$ is the set of elements in M_{k+1} but not in M_k for $k = 0, 1, 2, \dots$.

More explicitly, we shall show that every M_k can be picked as a field $M_k = S_{e_k}$ from the Steinitz tower (§4, (1)). In other words, we shall exhibit integers $0 = e_0 < e_1 < e_2 < \cdots$ such that the conditions (i) and (ii) above obtain. That such fields S_{e_k} form a tower (1) is trivial, while (i) follows from Theorem 4.2. To establish (ii), we shall show by induction that if the integers $e_0 < e_1 < e_2 < \cdots < e_k$ have already been chosen, there is an integer $e_{k+1} > e_k$ such that $M_k = S_{e_k} \subset L(D_{k+1}^p)$. Here

$$D_{k+1} = S_{e_{k+1}} - S_{e_k} \supset S_{e_{k+1}} - S_{e_{k+1}-1} = C_{e_{k+1}},$$

where $C_n = S_n - S_{n-1}$, as in Lemma 4.3. Hence it will suffice to demonstrate $S_{e_k} \subset L(C_{e_{k+1}}^p)$. This is a consequence of the following lemma:

LEMMA 5.2. *For any integer $e \geq 0$, there exists an integer $m > e$ so that in the Steinitz tower $S_e \subset L(C_m^p)$ where $C_m = S_m - S_{m-1}$.*

The proof will depend essentially upon the finiteness of T and the "relative perfection" of K over L . This latter property we assume in the form (cf. Theorem 3.5) $K = K^p(L) = L(K^p)$. Let T_e be a separating transcendence basis, obtained as in Theorem 4.2, for S_e over L . The basis T_e is finite because T is, while $T_e \subset L(K^p)$; so there is a finite set R of elements of K such that $T_e \subset L(R^p)$. Since $K = \sum S_e$, each element of R is in some one Steinitz field S_e , so that there is a finite integer $m > e$ such that $R \subset S_m$. Combining these conclusions, we have $T_e \subset L(S_m^p)$.

Consider now any element z in S_e . By the construction of T_e , z is separable over $L(T_e)$ and hence over $L(S_m^p)$. But z is also in S_e , hence in S_m since $m > e$. Therefore z^p is in $L(S_m^p)$, so that z is also purely inseparable over $L(S_m^p)$. This implies that z is in $L(S_m^p)$, so that Lemma 4.3 gives

$$(2) \quad S_e \subset L(S_m^p) = L(C_m^p)$$

as required for the lemma.

Theorem 5.1 is now established under the essential hypothesis that the transcendence basis is finite. Examples readily show that the same method cannot be used when T is infinite. However T will certainly be finite when the transcendence degree of K over its subfield P is finite. This special case we reformulate as follows:

THEOREM 5.3. *If K has a finite transcendence degree over a perfect subfield P , then there exists a tower of subfields*

$$(3) \quad L \subset M_0 \subset M_1 \subset M_2 \subset \cdots \subset K, \quad K = \sum_{k=0}^{\infty} M_k,$$

all containing P , such that

- (i) L has a separating transcendence basis over P ;
- (ii) Each field M_k has a separating transcendence basis over L ;
- (iii) L is relatively algebraically closed in K , and $K = K^p(L)$;
- (iv) $M_k \subset L(D_{k+1}^p)$ where $D_{k+1} = M_{k+1} - M_k$, for $k = 0, 1, \dots$.

6. **Exponent lemmas.** The difficulties in the way of proving properties (i) and (ii) of Schmidt's theorem for arbitrary Steinitz towers will be subsequently illustrated by elaborate examples, which require as a preliminary the structure of the Steinitz tower for a purely transcendental extension of a perfect field.

LEMMA 6.1. *If T is a set of elements algebraically independent over the field F , and if $K = F(T^{p^{-\infty}})$ is the field obtained from F by the adjunction of all elements*

$t^{p^{-e}}$ for e any integer and t in T , then for any e

$$(1) \quad \mathfrak{S}_e(K; F(T)) = F(T^{p^{-e}}).$$

In other words, $F(T^{p^{-e}})$ is exactly the set of the elements α of K such that α^{p^e} is separable over $F(T)$. A Steinitz field tower for K over $F(T)$ is then

$$(2) \quad F(T) \subset F(T^{p^{-1}}) \subset F(T^{p^{-2}}) \subset \dots$$

Proof. $S_e = \mathfrak{S}_e(K; F(T))$ denotes the set of all α in K such that α^{p^e} is separable over $F(T)$. That $S_e \supset F(T^{p^{-e}})$ results immediately, so that we need only prove the converse $F(T^{p^{-e}}) \supset S_e$. Since any element of S_e depends algebraically on but a finite number of the elements of T , it suffices to give a proof for the case when T is finite. We treat this case by an induction on the number of elements in T .

Case 1. T has one element t . Over the field S_e any power $z = t^{p^{-m}}$, with $m > e$ satisfies an equation $z^{p^{m-e}} = t^{p^{-e}}$. This equation is irreducible over S_e because otherwise the p th root of $t^{p^{-e}}$ is in S_e . This would imply that $t^{p^{-1}}$ is in S_0 and hence is separable as well as purely inseparable over $F(t)$. Consequently, $t^{p^{-1}}$ is in $F(t)$, an impossibility. Therefore $z^{p^{m-e}} = t^{p^{-e}}$ is irreducible over S_e , and the degree of $z = t^{p^{-m}}$ is

$$(3) \quad [S_e(t^{p^{-m}}):S_e] = p^{m-e}, \quad m > e.$$

Suppose now that an element α of S_e is not in $F(t^{p^{-e}})$. For a sufficiently large m , $\alpha \notin F(t^{p^{-m}})$. As $S_e \supset F(t^{p^{-e}}, \alpha) > F(t^{p^{-e}})$, we have according to (3) the following degree relations:

$$[F(t^{p^{-m}}):F(t^{p^{-e}}, \alpha)] < [F(t^{p^{-m}}):F(t^{p^{-e}})] = p^{m-e},$$

$$[F(t^{p^{-m}}):F(t^{p^{-e}}, \alpha)] \geq [S_e(t^{p^{-m}}):S_e] = p^{m-e},$$

a contradiction. We have proven $\mathfrak{S}_e(K, F(t)) = F(t^{p^{-e}})$.

Case 2. Suppose next that the lemma is known when the transcendence basis has $n-1$ elements, and let $T = T_0 + \{t\}$ have n elements, so that T_0 has $n-1$ elements. K contains a subfield $F' = F(T_0^{p^{-\infty}})$ and $K = F'(t^{p^{-\infty}})$. Any α in K with α^{p^e} separable over $F(T)$ has α^{p^e} also separable over $F'(t)$ so that α is contained in $F'(t^{p^{-e}}) = F(t^{p^{-e}}, T_0^{p^{-\infty}})$ by the proof of the previous case. If we set $F_0 = F(t^{p^{-e}})$, then α is in $F_0(T_0^{p^{-\infty}})$ and has α^{p^e} separable over $F_0(T_0)$. Therefore, by the induction assumption, α is in

$$F_0(T_0^{p^{-e}}) = F(t^{p^{-e}}, T_0^{p^{-e}}) = F(T^{p^{-e}}),$$

as required in the assertion (1).

7. Irregular Steinitz field towers. We shall now show that the field tower M_k of Theorem 5.1 with the special property

$$M_k \subset L(D_{k+1}^p), \quad D_{k+1} = M_{k+1} - M_k,$$

can be taken to be a Steinitz field tower itself whenever the transcendence basis T has only one element, but not always in other cases.

THEOREM 7.1. *If the basis T of the hypothesis of §4 consists of exactly one element, then*

$$(1) \quad S_k = L(S_{k+1}^p), \quad k = 0, 1, 2, \dots$$

Proof. By Theorem 4.2 each field S_e has over L a s.t.b. consisting of one element t_e . Thus each S_e consists of the elements of K of exponent p or less over $L(t_{e-1})$. By reason of this symmetry it patently suffices to prove our conclusion $S_{e-1} = L(S_e^p)$ only for the case $e=1$. Since t_1 is in S_1 and not S_0 , it has exponent p over $L(t_0)$ and also over $P(X, t_0)$. By the exchange lemma, some element of $\{t_0\} + X$ can be exchanged with t_1^p . If t_0 is not so exchangeable, this means, as in the exchange lemma, that t_1^p is separable over $P(X, t_0^p)$, and that some x in X can be here exchanged with t_1^p . This exchange makes x separable over $P(X - \{x\}, t_0^p, t_1^p)$ and thus over $K^p(X - \{x\})$. Hence (Lemma III) x lies in $K^p(X - \{x\})$, counter to the p -independence of the set X .

It must then be possible to exchange t_0 with t_1^p . Hence t_0 is separable over $P(t_1^p, X) \subset L(t_1^p)$. By the transitivity of separability, every element of S_0 is then separable over $L(t_1^p) \subset L(S_1^p)$. Every element of S_0 is in S_1 and hence is also purely inseparable over $L(S_1^p)$. Combining these facts (Lemma III), we conclude that $S_0 \subset L(S_1^p)$, as required in the theorem.

We now show by an example that this theorem is not always true when T has more than one element. Over a perfect field P construct the field

$$(2) \quad K = P(x, y^{p^{-\infty}}, z^{p^{-\infty}}) = P(x, y, z, y^{p^{-1}}, z^{p^{-1}}, y^{p^{-2}}, z^{p^{-2}}, \dots)$$

where x, y , and z are algebraically independent over P . The element x is by inspection a p -basis of K (cf. Teichmüller [6, Theorem 18]). Furthermore $P(x)$ is relatively algebraically closed in $P(x, y^{p^{-\infty}}, z^{p^{-\infty}})$ because any field is relatively algebraically closed in a purely transcendental extension. $P(x)$ is then also relatively algebraically closed in K , so that the field L of our hypothesis (cf. §4), consisting of all elements algebraic over $P(x)$, here becomes $P(x)$ itself. If we now introduce the quantity

$$(3) \quad u = xz^{p^{-1}} + y^{p^{-1}}, \quad u^p = x^p z + y,$$

then $T = \{u, z\}$ is a transcendence basis for K over L , because $y = u^p - x^p z$. For the Steinitz field tower relative to this basis T , we shall demonstrate

$$(4) \quad S_0 = P(x, u, z), \quad S_1^p = P(x^p, y, z).$$

From these equations it is clear that $L(S_1^p) = P(x, y, z) < S_0$, unlike (1), so that here Theorem 5.1 certainly does not hold.

The first equation of (4) will be established if we show $S_0 \subset P(x, u, z)$. Since K is a purely inseparable extension of $L(y, z)$, K is also a purely inseparable extension of the larger field $P(x, u, z) \supset L(y, z)$, and any element of K is either in $P(x, u, z)$ or is purely inseparable over $P(x, u, z) = L(u, z)$. Therefore the field S_0 of separable elements of K is $P(x, y, z)$, as in (4).

The crux of the example is the second equation of (4). Note first that

$$(5) \quad S_1^p = S_0 \cap K^p.$$

Introduce the additional subfields

$$F = P(x^p), \quad B = P(x^p, y, z) = F(y, z),$$

so that the terms of (5) become

$$(6) \quad S_0 = B(x, u) = B((x^p)^{1/p}, (x^p z + y)^{1/p}), \quad K^p = B(y^{p^{-\infty}}, z^{p^{-\infty}}).$$

Any element α of the intersection $S_0 \cap K^p$ is in S_0 and so has a p th power α^p separable over $B = F(y, z)$, by (6). But y and z are algebraically independent over F , so that by Lemma 6.1, applied to α and to the field K^p , α must be in $F(y^{p^{-1}}, z^{p^{-1}})$. The expression (5) can then be rewritten as

$$(7) \quad S_1^p = S_0 \cap F(y^{p^{-1}}, z^{p^{-1}}) = B((x^p)^{1/p}, (x^p z + y)^{1/p}) \cap B(y^{1/p}, z^{1/p}).$$

The generators x^p , y , and z of B are algebraically independent over P . Under these conditions the intersection on the right of (7) has been shown to be B itself.* This establishes the second half of (4).

The field K of this counter-example has a relatively simple structure, for K is simply $P'(x)$ where $P' = P(y^{p^{-\infty}}, z^{p^{-\infty}})$ is the maximal perfect subfield of K . The field K has a s.t.b. over this field P' . The example, however, can be so modified that this simple alternative description of its structure is not possible. We now construct such a modification in which the base field P is itself the maximal perfect subfield of K .

Over the perfect field P , consider four denumerable sets of quantities

$$Y = \{y_1, y_2, \dots\}, Z = \{z_1, z_2, \dots\}, V = \{v_1, v_2, \dots\}, W = \{w_1, w_2, \dots\}.$$

Let the elements of the set $V + W + \{y_1, z_1\}$ be algebraically independent over P . Define the remaining elements by the equations

$$(8) \quad y_{k+1}^p = v_k + y_k, \quad z_{k+1}^p = w_k + z_k, \quad k = 1, 2, 3, \dots,$$

* Mac Lane [3, §6]. The intersection was computed to show that the lattice of all fields between B and $B^{1/p}$ is not a modular lattice in the sense of G. Birkhoff; that is, is not a Dedekind structure in the terminology of O. Ore.

and construct the field $K = P(V, W, Y, Z)$. Then one can show that $X = V + W$ is a p -basis of K , that the corresponding field L is $P(V, W)$, and that $T = \{u, z_1\}$ is a transcendence basis for K over L , where u is defined by $u = v_1 z_2 + y_2$. Relative to this transcendence basis, the Steinitz field tower begins with the field $S_0 = L(u, z_1)$. By an extension of the argument of the last example we compute $S_1^p = P(V^p, W^p, y_2^p, z_2^p)$ and hence find that $L(S_1^p) = P(V, W, y_1, z_1)$. This field does not contain the element u of S_0 , for

$$u^p = (v_1 z_2 + y_2)^p = v_1^p (w_1 + z_1) + v_1 + y_1$$

is an irreducible equation for u over $L(S_1^p)$. Hence S_0 is not contained in $L(S_1^p)$, and the conclusion of Theorem 7.1 does not hold for this example.

Furthermore, in this case the maximal perfect subfield K^{p^∞} of K is the base field P . For $P(Y, Z)$ contains all elements v_k and w_k by (8) and hence is the whole field K . Furthermore, the set $Y + Z$ is algebraically independent over P , and it can be readily seen* that the maximal perfect subfield of such a purely transcendental extension $K = P(Y, Z)$ is simply the base field P itself. In conclusion, we can state the theorem:

THEOREM 7.2. *If the field K of the hypothesis of §4 has the transcendence degree 2 or more over the intermediate field L , then the fields of the Steinitz towers do not always satisfy the condition $S_0 = L(S_1^p)$ of Theorem 7.1. Specifically, there exist such fields K with maximal perfect subfield P and $S_0 > L(S_1^p)$.*

8. Inseparable Steinitz field towers. If the field K under consideration does not have a finite transcendence degree over its subfield L , as assumed in the treatment of §4, then the fields S_k of the Steinitz field tower need not all have separating transcendence bases over L . This we shall show by an example (which is summarized below in Theorem 8.6).

Let P be any perfect field, and consider two denumerable sets of elements

$$T = \{t_0, t_1, t_2, \dots\}, \quad Y = \{y_2, y_3, \dots\};$$

let the elements of the set T be algebraically independent over P , define the elements y of Y by the equations

$$(1) \quad y_n^p = t_{n-2} + t_{n-1}t_n^p, \quad n = 2, 3, 4, \dots,$$

and take K to be the field

$$K = P(T, Y^{p^{-\infty}}) = P(T, Y, Y^{p^{-1}}, Y^{p^{-2}}, \dots).$$

LEMMA 8.1. *The set X composed of t_0 alone is a p -basis of K .*

* Added in proof: A proof is given in S. Mac Lane, *Modular fields, I, Separating transcendence bases*, Duke Mathematical Journal, vol. 5 (1939). See Theorem 19, Corollary 1.

Proof. The defining algebraic equations (1) can be rewritten as

$$(2) \quad t_n = (y_{n+1}^p - t_{n-1})/t_{n+1}^p, \quad n \geq 1.$$

An induction on n proves that each t_n is in $K^p(t_0)$, so that $T \subset K^p(t_0)$, $Y^{p^{-\infty}} \subset K^p$, and hence $K = K^p(t_0)$. This is the first condition that t_0 be a p -basis. On the other hand, t_0 is p -independent; that is to say, t_0 is not in K^p . For suppose that $t_0 \in K^p$. The $n+1$ algebraically independent elements t_0, t_1, \dots, t_n are algebraic over $P(t_0, t_1, y_2, \dots, y_n)$ by the equations (1). Consequently the $n+1$ elements $t_0, t_1, y_2, \dots, y_n$ must themselves be independent (algebraically) over P . Hence $Y + \{t_0, t_1\}$ is a set of elements independent over P , and $\{t_0, t_1\}$ are likewise independent over the subfield $P(Y^{p^{-\infty}})$ of K . Introduce the additional subfields $K_n = P(Y^{p^{-\infty}}, t_0, t_1, \dots, t_n)$, with $K = \sum K_n$. By the equations (1) the t 's in this field K_n can be expressed rationally in terms of the y 's and the last two t 's. Hence $K_n = P(Y^{p^{-\infty}}, t_{n-1}, t_n)$, and $\{t_{n-1}, t_n\}$ is a set algebraically independent over $P(Y^{p^{-\infty}})$. Suppose now that t_0 is in K^p . Since $K = \sum K_n$, t_0 is in some field

$$K_n^p = P(Y^{p^{-\infty}}, t_{n-1}^p, t_n^p) = P(Y^{p^{-\infty}}, t_0^p, t_1^p, \dots, t_n^p).$$

A successive application of the equations (2) then shows that t_1, t_2, \dots , and finally t_{n-1} are also in K_n^p . But the elements t_{n-1}^p, t_n^p are known to be algebraically independent over $P(Y^{p^{-\infty}})$, so that the extended field $K_n^p = P(Y^{p^{-\infty}}, t_{n-1}^p, t_n^p)$ certainly cannot contain a p th root $t_{n-1} = (t_{n-1}^p)^{1/p}$. This contradiction shows that t_0 is p -independent.

LEMMA 8.2. *The field L of all elements algebraic over $P(t_0)$ is $L = P(t_0)$.*

Proof. By Theorem 3.5, any element α in L is separable algebraic over $P(t_0)$ and so over $P(T)$. But K is obtained from $P(T)$ by the successive adjunction of p th roots, which means that K is purely inseparable over $P(T)$. Therefore (Lemma III) the elements α of L all lie in $P(T)$. The remaining elements of T are algebraically independent of t_0 ; so L must be $P(t_0)$, as asserted.

We now choose for K over L the transcendence basis $T_1 = \{t_1, t_2, \dots\}$.

LEMMA 8.3. *The Steinitz field $S_1 = \mathfrak{S}(K; L(T_1))$ of all elements of K of exponent p or less over $L(T_1)$ is the field $S_1 = L(T_1, Y) = P(t_0, T_1, Y)$.*

The defining equations (1) for the elements y make each y of exponent p over $P(t_0, T_1)$. Hence $L(T_1, Y) \subset S_1$. Conversely, S_1 consists of certain elements of $K^{p^{-\infty}} = P(T^{p^{-\infty}})$ of exponent p or less over $P(T)$. Therefore, by Lemma 6.1, $S_1 \subset P(T^{p^{-1}})$. In other words, S_1 satisfies

$$(3) \quad M_1 \subset S_1 \subset M_2, \quad M_1 = L(T_1, Y), \quad M_2 = P(T^{p^{-1}}).$$

The equations (2) show that M_2 can also be generated as

$$M_2 = P(t_0^{p^{-1}}, T_1^{p^{-1}}) = P(T, Y, t_0^{p^{-1}}) = M_1(t_0^{p^{-1}}).$$

Therefore the field M_2 of (3) has degree p or 1 over M_1 , so that S_1 is necessarily M_1 or M_2 . If $S_1 = M_2$, then $t_0^{p^{-1}}$ is in $S_1 \subset K$; hence t_0 is in K^p , contrary to the result of Lemma 8.1. Therefore $S_1 = M_1 = L(T_1, Y)$.

LEMMA 8.4. *The field S_1^p contains neither t_n nor t_n/t_{n+1} for any integer $n \geq 0$.*

Proof. If t_n were in S_1^p , the equation (2) solved for t_{n-1} shows that t_{n-1} is in S_1^p . A repetition of this argument shows that t_{n-2} , t_{n-3} , and finally t_0 are in $S_1^p \subset K^p$, in contradiction to Lemma 8.1.

On the other hand, if t_n/t_{n+1} is in S_1^p , the equation (1) written in the form $y_{n+2}^p/t_{n+1} = t_n/t_{n+1} + t_{n+2}^p$ would imply that $1/t_{n+1}$ and hence t_{n+1} are in S_1^p , contrary to the already established part of the lemma.

LEMMA 8.5. *The first field $S_1 = L(T_1, Y)$ of the Steinitz tower does not have a separating transcendence basis over L .*

If there were such a basis over $L = P(t_0)$, the adjunction of t_0 to this basis would yield an enlarged s.t.b. $Z = \{z_1, z_2, \dots\}$ for S_1 over P . We shall show that this leads to a contradiction by finding a single z the adjunction of which would simultaneously make y_k and t_k separable, in conflict with the form of the inseparable defining equation (1). The argument depends on a reduction to a finite subset of Z . Specifically, both t_0 and t_1 are separable over $P(Z)$, so that there must be a finite subset $Z_m = \{z_1, \dots, z_m\}$ so large that t_0 and t_1 are separable over $P(Z_m)$. All of the independent elements of T cannot be dependent on this subset Z_m , so that there must be an integer $n \geq 2$, such that t_0, t_1, \dots, t_{n-1} are algebraic and hence separable over $P(Z_m)$, while the next element t_n is not so algebraic over Z_m . However, t_n will be algebraic over a larger set of z 's, so that there is a set $Z_k = \{z_1, \dots, z_k\}$, ($k \geq m$), for which t_n is algebraic over $P(Z_k, z_{k+1})$, but not over $P(Z_k)$. The equations (1) make (a) y_n algebraic over $P(t_{n-2}, t_{n-1}, t_n)$, (b) t_n algebraic over $P(t_{n-2}, t_{n-1}, y_n)$. Since both t_{n-2} and t_{n-1} are already algebraic over $P(Z_k) \supset P(Z_m)$, neither t_n nor y_n can be algebraic over $P(Z_k)$, and both t_n and y_n must be algebraic over $P(Z_k, z)$, where $z = z_{k+1}$.

From the equations for t_n and y_n over $P(Z_k, z)$, we can, by Lemma II, pick the largest integers e and f such that

(4) t_n is separable over $P(Z_k, z^{p^e})$; y_n is separable over $P(Z_k, z^{p^f})$.

By the exchange lemma, we then have

(5) z^{p^e} separable over $P(Z_k, t_n)$; z^{p^f} separable over $P(Z_k, y_n)$.

If $e \geq f$, the first statement of (4) and the second statement of (5) imply that

t_n is separable over $P(Z_k, y_n)$. Let N denote the field of all elements of S_1 separable over $P(Z_k)$. By construction, t_{n-2} and t_{n-1} are in N so that (1) makes t_n purely inseparable over $N(y_n)$. Therefore $t_n \notin N(y_n)$. In other words, t_n is a rational function

$$t_n = f(y_n)/g(y_n), \quad f(y_n), g(y_n) \text{ in } N[y_n],$$

where we can assume that the coefficients $f(0)$, $g(0)$ are not both 0. This value of t_n substituted in (1) yields

$$y_n^p [g(y_n)]^p = t_{n-2} [g(y_n)]^p + t_{n-1} [f(y_n)]^p.$$

Here the variable y_n over N can be replaced by 0 with the result

$$-t_{n-2} [g(0)]^p = t_{n-1} [f(0)]^p.$$

One and consequently both of $f(0)$, $g(0)$ are different from 0. Therefore $t_{n-2}/t_{n-1} = -[f(0)/g(0)]^p$ is in S_1^p , in contradiction to Lemma 8.4.

In the remaining case, when $e < f$, a similar argument proves $y_n \notin N(t_n)$ and hence $t_{n-2} \notin N^p \subset K^p$, another contradiction. We have therefore constructed a Steinitz field tower in which one of the fields S_1 has no s.t.b. over the ground field L .

THEOREM 8.6. *There is a modular field K with maximal perfect subfield P , a p -basis X , and a transcendence basis T over the subfield L of elements algebraic over $P(X)$, such that some field of the Steinitz tower for K relative to T over L does not have a separating transcendence basis over L .*

The example given establishes this theorem except for the hypothesis that P is the maximal perfect subfield of K ; for the maximal perfect subfield of the field used above manifestly includes $P(Y^{p^{-\infty}})$. The following modification of the example will complete this point.

Choose sets of elements

$$T = \{t_k\}, \quad X = \{x_{ij}\}, \quad Y = \{y_{ij}\}, \\ k = 0, 1, 2, \dots; i = 2, 3, 4, \dots; j = 0, 1, 2, \dots,$$

where the elements of $T+X$ are to be viewed as algebraically independent over a perfect field P , and where the elements y_{ij} are algebraic over $P(T, X)$ in accord with the equations

$$(6) \quad y_{i0}^p = t_{i-2} + t_{i-1} t_i^p, \quad i = 2, 3, \dots,$$

$$(7) \quad y_{ij+1}^p = x_{ij} + y_{ij}, \quad i = 2, 3, \dots; j = 0, 1, 2, \dots$$

Equations (6) are analogous to the defining equations (1) of the previous ex-

ample, while equations (7) differ from the repeated p th roots $Y^{p^{-n}}$ of the previous example only in the presence of the x_{ij} , which will insure that P is the maximal perfect subfield. The field K to be considered is $K = P(T, X, Y)$.

LEMMA 8.61. (Compare Lemma 8.1.) *The set $X + \{t_0\}$ is a p -basis of K .*

That $K = K^p(X, t_0)$, one sees by inspection of the equations (6) and (7). Conversely, to prove the p -independence of $X + \{t_0\}$ it suffices to prove that each $X_n + \{t_0\}$ is p -independent, where X_n is the first of the "truncated" sets

$$X_n = \{x_{ij}\}, \quad Y_n = \{y_{ij}\}, \quad i = 2, \dots, n; j = 0, \dots, n.$$

The field K is approximated by a tower of fields

$$(8) \quad K_n = P(t_0, t_1, \dots, t_n, X_n, Y_n).$$

Since any p -dependence will occur at some stage in this tower, it will suffice to prove $X_n + \{t_0\}$ p -independent in K_n . The equations (7) allow rational computations of y_{ij} with $j < n$ in terms of y_{in} , while according to (6), t_2, \dots, t_n are algebraic over $Y_n + \{t_0, t_1\}$. Hence K_n has the transcendence basis

$$U = X_n + \{t_0, t_1, y_{2n}, \dots, y_{nn}\}.$$

Specifically, over $P(U)$, K_n is the algebraic extension $K_n = P(U, t_2, \dots, t_n)$, of degree $[K_n : P(U)] \leq p^{n-1}$. $P(U)$ has a p -basis of $m+n+1$ elements, where m is the number of elements in X_n , so that K_n , as a finite purely inseparable extension of $P(U)$, has a p -basis of the same number* of elements. By the definition of p -independence in terms of degrees this means that $[K_n : K_n^p] = p^{m+n+1}$. Hence

$$(9) \quad [K_n : K_n^p(t_0, X_n)] \cdot [K_n^p(t_0, X_n) : K_n^p] = p^{m+n+1}.$$

On the other hand, $F_n = K_n^p(t_0, X_n)$ contains all p th powers from K_n , while by repeated applications of (6) it must contain t_1, t_2, \dots, t_{n-1} . But K_n is generated over P by X_n, t_0, \dots, t_n and y_{2n}, \dots, y_{nn} , so that

$$K_n = [K_n^p(t_0, X_n)](t_n, y_{2n}, \dots, y_{nn}).$$

Each element adjoined on the right is purely inseparable of exponent p or 1; hence $[K_n : K_n^p(t_0, X_n)] \leq p^n$. Combined with (9), this yields the inequality $[K_n^p(t_0, X_n) : K_n^p] \geq p^{m+1}$, where $m+1$ is the number of elements in $X_n + \{t_0\}$. Therefore $X_n + \{t_0\}$ is p -independent in K_n , as required for Lemma 8.61.

Using this p -basis, denote by L the field of those elements of K algebraic over $P(t_0, X)$, and consider the transcendence basis $T_1 = \{t_1, t_2, \dots\}$ for K over L .

* By a theorem (unpublished) due to Dr. M. Becker, or by direct computation in this case.

LEMMA 8.62. *The first field $S_1 = \mathfrak{S}_1(K; L(T_1))$ of the Steinitz tower relative to $L(T_1)$ is $S_1 = P(T, X, y_{20}, y_{30}, y_{40}, \dots)$.*

Proof. That S_1 includes the quantities indicated is manifest from the defining equations; so the conclusion could be false only in the presence of an element w not in $P(T, X, y_{20}, \dots)$ but in S_1 . The p th power w^p is then separable over $L(T_1)$ and hence over $P(t_0, X, T_1)$, by Theorem 3.5. Choose n so that w is in K_n of (8) and so that w^p is separable over the field

$$(10) \quad D_n = P(t_0, t_1, \dots, t_n, X_n).$$

The defining equations (7) for y_{in} can be combined as

$$(11) \quad y_{in}^{p^{n+1}} = x_{in-1}^{p^n} + x_{in-2}^{p^{n-1}} + \dots + x_{i0}^p + t_{i-1}t_i^p + t_{i-2}.$$

These equations have the form $y_{in}^{p^{n+1}} = u_{i-2}$, where the quantities u on the right lie in D_n and can be successively exchanged with the corresponding t_{i-2} in (10) to yield the generation

$$D_n = P(u_0, u_1, \dots, u_{n-2}, t_{n-1}, t_n, X_n).$$

The field $K_n = P(U, t_2, \dots, t_n)$ of (8) becomes

$$K_n = P(X_n, t_0, t_1, \dots, t_n, y_{2n}, \dots, y_{nn}),$$

and hence is generated by adjoining to D_n the roots $y_{in} = u_{i-2}^{p^{-n-1}}$:

$$K_n = D_n(u_0^{p^{-n-1}}, u_1^{p^{-n-1}}, \dots, u_{n-2}^{p^{-n-1}}).$$

The element w of K_n of exponent 1 over D_n must then by Lemma 6.1 (applied with $F = P(t_{n-1}, t_n, X_n)$) lie in the field $D_n(u_0^{1/p}, u_1^{1/p}, \dots, u_{n-2}^{1/p})$. By the expansions for the u 's on the right of (11), this is the field $D_n(y_{20}, y_{30}, \dots, y_{n0})$. This field is contained in the field $P(T, X, y_{20}, y_{30}, \dots)$ of the lemma, counter to the assumption that w does not lie in this field. This field is therefore equal to S_1 , as asserted in the lemma.

This field S_1 may be briefly described as the field $S_1 = P(T, X, y_{20}, y_{30}, \dots)$ generated by the adjunction of the independent variables T, X , and the roots y_{i0} of the equations (6). It differs from the field S_1 of the previous example only in the presence of certain variables X which nowhere figure in the defining equations (6). A reapplication of the arguments used in the previous case (Lemmas 8.1, 8.2, and 8.5) then establishes the following lemma:

LEMMA 8.63. *The first field S_1 of the Steinitz tower does not have a separating transcendence basis over L .*

This completes the counter-example, with the following additional property not present in the previous example:

LEMMA 8.64. *The field K above has P as its maximal perfect subfield.*

Proof. Embed K in the field $K' = K(s_0, s_1, \dots)$, where $s_i = t_i^{1/p}$. If Y' is the set of elements y_{ij} with $i=2, 3, \dots$ and $j=1, 2, \dots$, then $K' = P(Y', s_0, s_1, \dots)$, by the defining equations (6). Furthermore, the generators $Y' + \{s_0, s_1, \dots\}$ are algebraically independent. For the set of elements $\{t_0, t_1, \dots, t_m, x_{ij}\}$ with $i=2, \dots, m$ and $j=0, \dots, n-1$ consists of $(m+1) + (m-1)n$ elements and is known to be algebraically independent, but is algebraically dependent upon the set $\{s_0, \dots, s_m, y_{ij}\}$ with $i=2, \dots, m$ and $j=1, \dots, n$, which has the same number of elements. Therefore this subset and the whole set $Y' + \{s_0, s_1, \dots\}$ are algebraically independent. The purely transcendental field $K' = P(Y', s_0, s_1, \dots)$ therefore has P as maximal perfect subfield, as asserted.

9. Separating linear orders of the Steinitz field tower. F. K. Schmidt has considered the possibility of "separating" orders for fields. Let a set K which is a field have a linear order given by a relation $<$. For any element b in K let K_b denote the subfield of K generated by the set of all elements c with $c < b$. The given linear order is said to be a *separating order* if every element b of K is either transcendental or separable and algebraic over the corresponding K_b . The elements b algebraic over their respective fields K_b are said to be *algebraic* in the given order. Schmidt [2, pp. 16, 46] now considers the following situation: * K is a field which has no separating transcendence basis over its prime field P ; L is a subfield of K with a separating transcendence basis over P such that $K^{p^{-\infty}} = K(L^{p^{-\infty}})$; T is any transcendence basis for K over L , and S_n is again the Steinitz field composed of all elements α of K such that α^{p^n} is separable over $L(T)$. This situation includes, in particular,† the situation described in the hypothesis in our §4, provided we suppose that the P used there is the prime field $GF[p]$ and that K has no separating transcendence basis over P . (Both of these assumptions can be made in the examples of fields constructed in §§7 and 8.)

Given any such situation, Schmidt now asserts *without proof* [2, p. 46] that "there exists a separating order ' $<$ ' of K such that (i) ' $<$ ' induces in each field S_n a separating normal order W_n (well ordering); (ii) if the elements of K are written down in the order specified by ' $<$ ', then one obtains an additive representation

* The notation has been changed thus:

F. K. Schmidt:	\mathfrak{K}	\mathfrak{L}	\mathfrak{S}	\mathfrak{S}_i	\mathfrak{D}
S. Mac Lane:	K	L	T	S_i	$<$

† Schmidt does not assume that his field L can be constructed from a p -basis X ; examples can be given of a field L which cannot be so constructed and which still has the properties specified by Schmidt.

$$K = L + \sum_{n=0}^{\infty} C_n, \quad C_n = S_n - S_{n-1}, \quad C_0 = S_0 - L.$$

In other words, the elements of L precede all other elements, and the elements of each complement C_n preceded those of the complement C_{n-1}, \dots (iii) Every element b of S_n algebraic in the order ' $<$ ' of K is separable and algebraic in the separating normal order W_{n+1} of the subfield S_{n+1} . Furthermore, the coefficients of the irreducible separable polynomial $G(x)$ satisfied by b in the order W_{n+1} (that is, satisfied by b over $(S_{n+1})_b$) are present in the field $S_{nb}(C_{n+1}^p)$ where S_{nb} is the smallest subfield of S_n containing all elements of S_n which precede b in the order W_n .

The separating order W_n obtained here means that S_n contains no element inseparable and algebraic over the intermediate field L . In other words, K can contain no such element. The hypotheses stated for L are not in themselves sufficient to insure this condition. Certainly an additional hypothesis is intended, such as the assumption that L is relatively algebraically closed in K or the assumption that the elements of K algebraic over L are separable over L .

The conclusion (iii) formulated above can be further reduced. For any b in S_n the field $(S_{n+1})_b$ which contains all elements of S_{n+1} preceding b must by (ii) contain L and C_{n+1} , and, therefore, by Lemma 4.3, also contains $L(C_{n+1}) = L(S_{n+1}) = S_{n+1}$. In other words, b is contained in $(S_{n+1})_b$; the irreducible equation $G(x)$ is $x - b$, and condition (iii) becomes

$$(1) \quad b \in S_{nb}(C_{n+1}^p).$$

We now show from (iii) by transfinite induction that every b of S_n is in $L(C_{n+1}^p)$. The first b of S_n lies, by condition (ii), in L and hence in $L(C_{n+1}^p)$. Suppose now that our assertion has been established for all predecessors of b in the normal order of W_n of S_n . The field S_{nb} is then generated by elements $c < b$ which, by assumption, are all in $L(C_{n+1}^p)$; hence by (1), b is also in $L(C_{n+1}^p)$. Since K is supposed to have no separating transcendence basis, $K > L$ and Lemma 4.3 applies. It shows that $L(C_{n+1}^p) = L(S_{n+1})$, so that the conclusion obtained can be stated thus:

LEMMA 9.1. *Conditions (i), (ii), and (iii) above imply that $S_n \subset L(S_{n+1}^p)$.*

This conclusion cannot always be true, as indicated in Theorem 7.2. Therefore the conclusion (iii) must be dropped. On the other hand, (i) means, as in Theorem 2.1, that each S_n has a separating transcendence basis over P . That this cannot always be the case was shown in Theorem 8.1. Schmidt's conclusions can then only be taken in some restricted form, as in our Theo-

rems 4.2 and 5.1, or perhaps by stating that for a field K there exists a specifically selected field L and transcendence basis T for which the conclusions are true. A restricted theorem of this latter type, if demonstrable, would be satisfactory for the applications to the structure of perfect fields envisaged by Schmidt.

BIBLIOGRAPHY

1. A. A. Albert, *Modern Higher Algebra*, Chicago, 1937.
2. H. Hasse and F. K. Schmidt, *Die Struktur diskret bewerteter Körper*, Journal für die reine und angewandte Mathematik, vol. 170 (1934), pp. 4-63.
3. S. Mac Lane, *A lattice formulation for transcendence degrees and p -bases*, Duke Mathematical Journal, vol. 4 (1938), pp. 456-468.
4. ———, *Subfields and automorphism groups of p -adic fields*, Annals of Mathematics, vol. 40 (1939), pp. 423-442.
5. E. Steinitz, *Algebraische Theorie der Körper*, edited by R. Baer and H. Hasse, Berlin, 1930.
6. O. Teichmüller, *p -Algebren*, Deutsche Mathematik, vol. 1 (1936), pp. 362-388.
7. ———, *Diskret bewertete perfekte Körper mit unvollkommenem Restklassenkörper*, Journal für die reine und angewandte Mathematik, vol. 176 (1937), pp. 141-152.
8. E. Witt, *Zyklische Körper und Algebren der Charakteristik p vom Grad p^n* , Journal für die reine und angewandte Mathematik, vol. 176 (1937), pp. 126-140.

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.
HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

ON INTERPOLATION BY FUNCTIONS ANALYTIC AND BOUNDED IN A GIVEN REGION*

BY
J. L. WALSH

The writer has recently formulated† the following problem, but without proving in detail any results on convergence of the sequences involved:

PROBLEM A. *Let the points $\beta_{n1}, \beta_{n2}, \dots, \beta_{nn}$, not necessarily distinct, lie interior to the region R of the plane of the complex variable z . Let the function $f(z)$ be analytic in each point β_{nk} . Let $f_n(z)$ be the (or a) function which coincides with $f(z)$ in the points $\beta_{n1}, \beta_{n2}, \dots, \beta_{nn}$, which is analytic in R , and the least upper bound M_n of whose modulus in R is a minimum. To study the functions $f_n(z)$, especially the approach to $f(z)$ of the sequence $f_n(z)$, and study the sequence M_n as n becomes infinite.*

A function $f_n(z)$ always exists (loc. cit.), and is unique if R is simply-connected.

It is the object of the present note to establish some results concerning Problem A, especially

THEOREM 1. *Let R be the interior of a Jordan curve C_1 . Let each of the points β_{nk} lie on or interior to a Jordan curve C_2 interior to C_1 , and let us suppose the relation*

$$(1) \quad \lim_{n \rightarrow \infty} |(z - \beta_{n1})(z - \beta_{n2}) \cdots (z - \beta_{nn})|^{1/n} = e^{V_1(x, y)}, \quad z = x + iy,$$

to hold at every point z exterior to C_2 , uniformly on any closed bounded set exterior to C_2 . Let $V_2(x, y)$ denote the function which coincides with $V_1(x, y)$ on C_1 and is harmonic interior to C_1 , continuous in the corresponding closed region. Let us suppose the function $V(x, y) = V_1(x, y) - V_2(x, y)$ to be continuous in the closure \bar{S} of the annular region S bounded by C_1 and C_2 , and to take the constant value γ at every point of C_2 . We denote generically by C_λ the locus $V(x, y) = \lambda$, ($\gamma < \lambda < 0$), in R , so that C_λ is a Jordan curve separating C_1 and C_2 ; we denote by R_λ the interior of C_λ , and by \bar{R}_λ the closed interior of C_λ .

Let the function $f(z)$ be analytic throughout the interior of R , but not throughout the interior of any $R_{\rho'}$, ($\rho' > \rho$). In the notation of Problem A, the sequence $f_n(z)$ converges uniformly to $f(z)$ on any closed set interior to R_ρ . Moreover we have ($\gamma < \sigma < \rho$)

* Presented to the Society, April 8, 1939; received by the editors October 3, 1938.

† Proceedings of the National Academy of Sciences, vol. 24 (1938), pp. 477-486.

$$(2) \quad \limsup_{n \rightarrow \infty} [\max |f(z) - f_n(z)|, z \text{ on } C_\sigma]^{1/n} = e^{\sigma-\rho},$$

$$(3) \quad \limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_n(z)|, z \text{ in } R]^{1/n} = e^{-\rho}.$$

1. **Proof of Theorem 1.** The technique of our study of Problem A is quite similar to the technique developed in a recent work* by the present writer, to which we shall make frequent reference.

The mere existence of the limit in (1) in R exterior to C_2 implies the uniformity of the limit on any closed bounded set exterior to C_2 (compare op. cit., p. 266). The function

$$(4) \quad U_n(x, y) = \frac{1}{n} \log |(z - \beta_{n1}) \cdots (z - \beta_{nn})|$$

is harmonic exterior to C_2 , so its limit $V_1(x, y)$ is also harmonic exterior to C_2 . Consequently the function $V(x, y)$ is harmonic in S .

If Γ is an analytic Jordan curve separating C_1 and C_2 , and if ν denotes the exterior normal for Γ , then the integral over Γ of $\partial U_n(x, y)/\partial \nu$ is 2π , whence (compare op. cit., p. 268)

$$(5) \quad 2\pi = \int_{\Gamma} \frac{\partial V_1}{\partial \nu} ds = \int_{\Gamma} \frac{\partial V}{\partial \nu} ds.$$

A consequence of (5) is the inequality $\gamma < 0$.

Let C'_1 be an analytic Jordan curve near C_1 containing C_1 in its interior. We shall eventually allow C'_1 to approach C_1 . Let $V'_2(x, y)$ denote the function which coincides with $V_1(x, y)$ on C'_1 and is harmonic interior to C'_1 , continuous in the corresponding closed region. The function

$$V'(x, y) = V_1(x, y) - V'_2(x, y)$$

is continuous in the closure \bar{S}' of the region S' bounded by C'_1 and C_2 and vanishes on C'_1 . As in the proof of (5) we have

$$(6) \quad 2\pi = \int_{C'_1} \frac{\partial V_1}{\partial \nu} ds = \int_{C'_1} \frac{\partial V'}{\partial \nu} ds.$$

As in the book cited, §9.11 (p. 265), we may write the following equations for (x, y) interior to C'_1 ; the second of these equations is a consequence of the corresponding equation with V_1 replaced by U_n :

* *Interpolation and Approximation by Rational Functions in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. 20, New York, 1935. All references in the present note not otherwise indicated are to this book, to which the reader is also referred for terminology.

$$\begin{aligned}
 V_2'(x, y) &= \frac{1}{2\pi} \int_{C_1'} \left(V_2' \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial V_2'}{\partial \nu} \right) ds, \\
 0 &= \frac{1}{2\pi} \int_{C_1'} \left(V_1 \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial V_1}{\partial \nu} \right) ds, \\
 V_2'(x, y) &= \frac{-1}{2\pi} \int_{C_1'} \left(V' \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial V'}{\partial \nu} \right) ds, \\
 (7) \quad V_2'(x, y) &= \frac{1}{2\pi} \int_{C_1'} \log r \frac{\partial V'}{\partial \nu} ds.
 \end{aligned}$$

The integrals are to be taken in the counterclockwise sense, and ν indicates the exterior normal.

When C_1' approaches C_1 , the function $V_2'(x, y)$ approaches $V_2(x, y)$ uniformly on and within C_1 , by Lebesgue's results on harmonic functions in variable regions.* Then the function $V'(x, y)$ approaches $V(x, y)$ uniformly in \bar{S} , and on C_2 the function $V'(x, y)$ takes on values uniformly as near as desired to $\gamma < 0$, provided merely that C_1' is sufficiently near to C_1 . Thus when C_1' is sufficiently close to C_1 , in S' we have $V'(x, y) < 0$ because $V'(x, y)$ is zero on C_1' and negative on C_2 , and on C_1' we have $\partial V'/\partial \nu \geq 0$; the equality sign is excluded here by our choice of C_1' as an *analytic* Jordan curve.

Let now the points $\alpha_{n1}, \alpha_{n2}, \dots, \alpha_{n, n-1}$ be chosen uniformly distributed on C_1' with respect to the parameter whose differential is the positive quantity $(\partial V'/\partial \nu)ds$ (compare op. cit., §§8.7 and 9.11). From (6) and (7) we have

$$\lim_{n \rightarrow \infty} |(z - \alpha_{n1})(z - \alpha_{n2}) \cdots (z - \alpha_{n, n-1})|^{1/n} = e^{V_2'(x, y)},$$

uniformly on any closed set interior to C_1' ; so by virtue of (1) we may write

$$(8) \quad \lim_{n \rightarrow \infty} \left| \frac{(z - \beta_{n1}) \cdots (z - \beta_{nn})}{(z - \alpha_{n1}) \cdots (z - \alpha_{n, n-1})} \right|^{1/n} = e^{V'(x, y)},$$

uniformly on any closed set interior to S' .

We denote by $r_n(z)$ the rational function of degree $n-1$ whose poles lie in the points $\alpha_{n1}, \alpha_{n2}, \dots, \alpha_{n, n-1}$ and which interpolates to $f(z)$ in each of the points $\beta_{n1}, \beta_{n2}, \dots, \beta_{nn}$; the sequence $r_n(z)$ has been studied in some detail (op. cit., §8.3), and in particular there can be established† the formula

$$(9) \quad \limsup [\max |r_n(z)|, z \text{ on } C_\mu']^{1/n} \leq e^{\mu - \rho'},$$

* Rendiconti del Circolo Matematico di Palermo, vol. 24 (1907), pp. 371-402.

† Inequality (9) is an immediate consequence of equation (8) and the standard formula for $r_n(z)$ (op. cit., p. 186), which is valid even exterior to C_ρ . Indeed, the sign \leq in (9) can be replaced by the equality sign, as the writer expects to indicate in a forthcoming paper in these Transactions.

where C'_λ denotes generically the Jordan curve $V'(x, y) = \lambda$ in S' , where $f(z)$ is analytic interior to C'_ρ , but is not analytic throughout the interior of any $C'_{\rho''}$, ($\rho'' > \rho'$), and where $\mu > \rho'$.

When C'_1 approaches C_1 , the locus C'_λ approaches uniformly the locus C_λ . Given any $\epsilon > 0$, we can choose C'_1 so near to C_1 that $|V'(x, y) - V(x, y)| < \epsilon$ uniformly in \bar{S} . For such a particular choice of C'_1 we have $\rho' > \rho - \epsilon$; the curve C_1 lies interior to some C'_μ , whence from (9)

$$(10) \quad \limsup_{n \rightarrow \infty} [\max |r_n(z)|, z \text{ on } C_1]^{1/n} \leq e^{\mu - \rho + \epsilon} \leq e^{-\rho + \epsilon}.$$

We have now exhibited functions $r_n(z)$ analytic in R , interpolating to $f(z)$ in the points β_{nk} , and satisfying (10). For the functions $f_n(z)$ whose least upper bound in R is a minimum we consequently have by (10)

$$(11) \quad \limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_n(z)|, z \text{ in } R]^{1/n} \leq e^{-\rho + \epsilon}.$$

A combination of (10) and (11) yields

$$\limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_n(z) - r_n(z)|, z \text{ in } R]^{1/n} \leq e^{-\rho + \epsilon},$$

whence for suitably chosen M ,

$$(12) \quad |f_n(z) - r_n(z)| \leq M e^{n(-\rho + 2\epsilon)}, \quad z \text{ in } R.$$

The function $f_n(z) - r_n(z)$ vanishes in each of the points β_{nk} ; so the familiar reasoning used in the proof of Schwarz's lemma gives, for z interior to C_1 ,

$$\begin{aligned} & \left| [f_n(z) - r_n(z)] \frac{(z - \alpha_{n1}) \cdots (z - \alpha_{n,n-1})}{(z - \beta_{n1}) \cdots (z - \beta_{nn})} \right| \\ & \leq M e^{n(-\rho + 2\epsilon)} / \left[\min \left| \frac{(z - \beta_{n1}) \cdots (z - \beta_{nn})}{(z - \alpha_{n1}) \cdots (z - \alpha_{n,n-1})} \right|, z \text{ on } C_1 \right]. \end{aligned}$$

For z on C_1 we have $V=0$, $V' > -\epsilon$; for z on C_σ , ($\gamma < \sigma < \rho'$), we have $V' < \sigma + \epsilon$; then by (8) we may write

$$(13) \quad \limsup_{n \rightarrow \infty} [\max |f_n(z) - r_n(z)|, z \text{ on } C_\sigma]^{1/n} \leq e^{\sigma - \rho + 4\epsilon}.$$

But we know also (op. cit., p. 198) for $\sigma' < \rho'$

$$\limsup_{n \rightarrow \infty} [\max |f(z) - r_n(z)|, z \text{ on } C_{\sigma'}]^{1/n} \leq e^{\sigma' - \rho'},$$

whence

$$(14) \quad \limsup_{n \rightarrow \infty} [\max |f(z) - r_n(z)|, z \text{ on } C_\sigma]^{1/n} \leq e^{\sigma - \rho + 2\epsilon}.$$

Inequalities (13) and (14) when combined now imply by letting ϵ approach zero ($\gamma < \sigma < \rho$)

$$(15) \quad \limsup_{n \rightarrow \infty} [\max |f(z) - f_n(z)|, z \text{ on } C_\sigma]^{1/n} \leq e^{\sigma - \rho}.$$

Likewise in (11) we may allow ϵ to approach zero:

$$(16) \quad \limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_n(z)|, z \text{ in } R]^{1/n} \leq e^{-\rho}.$$

To complete the proof of Theorem 1, it remains merely to show that the inequality sign cannot hold in (15) or (16). The proof is indirect; let us assume for instance

$$(17) \quad \limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_n(z)|, z \text{ in } R]^{1/n} \leq e^{-\rho_1}, \quad \rho_1 > \rho;$$

we shall reach a contradiction.

If $\eta > 0$ is arbitrary, we have from (15) for n sufficiently large

$$|f_{n+1}(z) - f_n(z)| \leq e^{(\sigma - \rho + \eta)n}, \quad z \text{ on } C_\sigma,$$

and we have from (17) for n sufficiently large

$$|f_{n+1}(z) - f_n(z)| \leq e^{(-\rho_1 + \eta)n}, \quad z \text{ in } R.$$

By an extension of Hadamard's Three-Circle Theorem* applied to the region bounded by C_1 and C_σ , we deduce for z on C_μ , ($\sigma < \mu < 0$),

$$(18) \quad |f_{n+1}(z) - f_n(z)| \leq [e^{(\sigma - \rho + \eta)n}]^{\mu/\sigma} \cdot [e^{(-\rho_1 + \eta)n}]^{(\sigma - \mu)/\sigma} = e^{(\mu\sigma + \eta\sigma - \mu\rho - \rho_1\sigma + \mu\rho_1)n/\sigma}.$$

Since σ is negative, the sequence $f_n(z)$ converges uniformly on C_μ provided merely

$$\phi(\mu) = \mu\sigma + \eta\sigma - \mu\rho - \rho_1\sigma + \mu\rho_1 > 0.$$

For the value $\mu = \rho$ the continuous function $\phi(\mu)$ takes the value

$$\phi(\rho) = \rho\sigma + \eta\sigma - \rho^2 - \rho_1\sigma + \rho\rho_1 = (\sigma - \rho)(\rho - \rho_1) + \eta\sigma.$$

By virtue of $\rho_1 > \rho$ and $\sigma < \rho$, it follows that when η is sufficiently small, $\phi(\rho)$ is positive. Consequently, $\phi(\mu)$ is positive also for suitably chosen values of μ greater than ρ . The limit of the sequence $f_n(z)$ is $f(z)$ interior to C_ρ , hence is the analytic function $f(z)$ throughout the interior of some curve C_μ , ($\mu > \rho$), which contradicts our definition of ρ .

* R. Nevanlinna, *Eindeutige Analytische Funktionen*, Berlin, 1936, p. 42. We are here using the Two-Constant Theorem (Zweikonstantensatz) in the form due to F. and R. Nevanlinna. A somewhat less precise form is due to Ostrowski. In the situation of Theorem 1 itself, but not in the more general situation described in §3, the Three-Circle Theorem can be applied after a conformal map by means of the function $w = \exp \{V(x, y) + iW(x, y)\}$, where $W(x, y)$ is conjugate to $V(x, y)$ interior to S .

We have now shown that the inequality sign in (16) is impossible. Precisely the same method shows that the inequality sign in (15) is impossible; so Theorem 1 is established.

A limiting case of (2) is also valid, namely

$$\limsup_{n \rightarrow \infty} [\max |f(z) - f_n(z)|, z \text{ on } C_2]^{1/n} = e^{\gamma - \rho};$$

indeed the obvious relation

$$\max [|f(z) - f_n(z)|, z \text{ on } C_2] \leq \max [|f(z) - f_n(z)|, z \text{ on } C_\sigma]$$

by approach of σ to γ establishes the precise analogue of (15), and the previous method shows the impossibility of the inequality.

2. Complements to Theorem 1. A complement to Theorem 1 is the

COROLLARY. *Under the conditions of Theorem 1 we have ($0 > \mu \geq \rho$)*

$$\limsup_{n \rightarrow \infty} [\max |f_n(z)|, z \text{ on } C_\mu]^{1/n} = e^{\mu - \rho}.$$

From (15) and (16) respectively we have ($\sigma < \rho$)

$$\limsup_{n \rightarrow \infty} [\max |f_{n+1}(z) - f_n(z)|, z \text{ on } C_\sigma]^{1/n} \leq e^{\sigma - \rho},$$

$$\limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_{n+1}(z) - f_n(z)|, z \text{ in } R]^{1/n} \leq e^{-\rho},$$

from which we deduce as in the proof of (18),

$$\limsup_{n \rightarrow \infty} [\max |f_{n+1}(z) - f_n(z)|, z \text{ on } C_\mu]^{1/n} \leq e^{\mu - \rho}.$$

We are now at liberty to write

$$\limsup_{n \rightarrow \infty} [\max |f_n(z)|, z \text{ on } C_\mu]^{1/n} \leq e^{\mu - \rho}.$$

The impossibility of the inequality sign here follows precisely as in (16) for $\mu > \rho$ and is trivial for $\mu = \rho$ (we should otherwise have $f_n(z)$ approaching zero uniformly interior to C_ρ); so the corollary is established.

It is of interest to note that when C_1 is a curve $V'(x, y) = \text{const.}$, it follows from (9) that the rational functions $r_n(z)$ have maximum modulus on C_μ , ($0 > \mu > \rho$), of the same order of magnitude as the maximum modulus of the extremal functions $f_n(z)$; a similar remark holds also of C_1 . Under these conditions it is likewise true that $\max |f(z) - r_n(z)|$ and $\max |f(z) - f_n(z)|$ have the same order of magnitude on C_σ , ($\rho > \sigma > \gamma$), and also on C_2 .

A relation which essentially includes (2) (granted the convergence of $f_n(z)$ to $f(z)$ in R_ρ) as well as the corollary, and thereby unifies the preceding results is

$$\limsup_{n \rightarrow \infty} [\max |f_{n+1}(z) - f_n(z)|, z \text{ on } C_\sigma]^{1/n} = e^{\sigma-\rho},$$

$0 > \sigma \geq \rho$ or $\rho > \sigma > \gamma$. This relation with the equality sign replaced by \leq has been pointed out in the proof of the corollary; if the inequality sign were to hold we should have the inequality sign in (2) or in the corollary, according as $\sigma < \rho$ or $\sigma \geq \rho$, which we know to be impossible. The corresponding limiting equations also hold and are similarly proved:

$$\limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_{n+1}(z) - f_n(z)|, z \text{ in } R]^{1/n} = e^{-\rho},$$

$$\limsup_{n \rightarrow \infty} [\max |f_{n+1}(z) - f_n(z)|, z \text{ on } C_2]^{1/n} = e^{\gamma-\rho}.$$

It is an obvious consequence of Theorem 1 that under the hypothesis of that theorem there exists no sequence of functions $F_n(z)$ analytic in R and coinciding with $f(z)$ in the points $\beta_{n1}, \beta_{n2}, \dots, \beta_{nn}$ such that we have

$$\limsup_{n \rightarrow \infty} [\text{l.u.b. } |F_n(z)|, z \text{ in } R]^{1/n} < e^{-\rho}.$$

We note too that Theorem 1 can be applied under the hypothesis of that theorem where C_μ , ($\mu > \rho$), plays the role of the original C_1 . The function $V(x, y) - \mu$ now takes the role of the original $V(x, y)$, and it follows from Theorem 1 that there exists a sequence of functions $F_n(z)$ analytic in R and coinciding with $f(z)$ in the points $\beta_{n1}, \beta_{n2}, \dots, \beta_{nn}$, namely the extremal functions $f_n(z)$ pertaining to R , such that we have

$$\limsup_{n \rightarrow \infty} [\text{l.u.b. } |F_n(z)|, z \text{ in } R_\mu]^{1/n} = e^{\mu-\rho};$$

but there exists no sequence of functions $F_n(z)$ analytic in R_μ and coinciding with $f(z)$ in the points $\beta_{n1}, \beta_{n2}, \dots, \beta_{nn}$ such that we have

$$\limsup_{n \rightarrow \infty} [\text{l.u.b. } |F_n(z)|, z \text{ in } R_\mu]^{1/n} < e^{\mu-\rho}.$$

Thus the extremal functions $f_n(z)$ of Theorem 1 have maximum moduli on C_μ , ($\mu > \rho$), which are of the same order of magnitude as the least upper bounds of the corresponding extremal functions which pertain to R_μ itself.

Still another remark is appropriate in connection with Theorem 1, relative to functions $f(z)$ analytic throughout R . Under these conditions we can set $\rho = 0$ in inequality (10), whence for the extremal functions $f_n(z)$ defined as in Theorem 1,

$$\limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_n(z)|, z \text{ in } R]^{1/n} \leq e^*.$$

Here we may allow ϵ to approach zero, whence

$$\limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_n(z)|, z \text{ in } R]^{1/n} \leq 1.$$

The inequality sign cannot hold here except in the trivial case $f(z) \equiv 0$, for the inequality sign implies that $f_n(z)$ approaches zero uniformly in R . As in the proof of (15) we have for every σ , ($\gamma < \sigma < 0$),

$$\limsup_{n \rightarrow \infty} [\max |f(z) - f_n(z)|, z \text{ on } C_\sigma]^{1/n} \leq e^\sigma.$$

If $f(z)$ is analytic and bounded in R , the sequence $f_n(z)$ is uniformly bounded in R , for $f(z)$ itself satisfies the conditions of interpolation:

$$[\text{l.u.b. } |f_n(z)|, z \text{ in } R] \leq [\text{l.u.b. } |f(z)|, z \text{ in } R].$$

There is evidence to indicate that the present methods alone do not enable us to determine the exact value of

$$\limsup_{n \rightarrow \infty} [\max |f(z) - f_n(z)|, z \text{ on } C_\sigma]^{1/n}, \quad 0 > \sigma > \gamma,$$

when $f(z)$ is analytic throughout R . First, there are various comparison sequences $r_n(z)$ any one of which is adequate in the proof of Theorem 1 itself but which yield different results for

$$\limsup [\max |f(z) - r_n(z)|, z \text{ in } R]^{1/n}$$

when $f(z)$ is analytic throughout R . This is shown for instance by choosing $f(z) = 1/(T-z)$, ($T > 1$), the β_{nk} as all zero, and the α_{nk} as the $(n-1)$ st roots of A^{n-1} , where $1 < A < T$, and by choosing $\beta_{nk} = 0$ and C_1 as $|z| = 1$. Equation (8) is fulfilled. The sequence $r_n(z)$ serves as a comparison sequence in the proof of Theorem 1 for an arbitrary function $f(z)$ satisfying the hypothesis of Theorem 1 without the necessity of allowing A to approach unity; that is to say, without the necessity of allowing C'_1 to approach C_1 : this is always the case when $V'(x, y)$ is constant on C_1 . It follows (as in op. cit., p. 185) that we have with the special choice of $f(z)$

$$f(z) - r_n(z) = z^n(T^{n-1} - A^{n-1})/[T^n(z^{n-1} - A^{n-1})(T - z)],$$

where $r_n(z)$ is found by interpolation to $f(z)$ in the points β_{nk} and has the poles α_{nk} . Consequently we may write

$$\limsup_{n \rightarrow \infty} [\max |f(z) - r_n(z)|, \text{ for } |z| = r \leq 1]^{1/n} = r/A,$$

whereas A is completely arbitrary within the limits $1 < A < T$, and its use is entirely accidental in the study of the functions $f_n(z)$.

Second, even when the singularities of the function $f(z)$ fall in the region in which (8) is valid, and when $V'(x, y)$ is constant on C_1 so that Theorem 1 itself can be established without varying the curve C_1' or the points α_{kn} , it is not true that the degree of convergence to $f(z)$ on C_r , ($0 < \sigma < \gamma$), is necessarily the same for the sequences $r_n(z), f_n(z)$. Let β be arbitrary, ($0 < \beta < 1$), and set

$$f(z) = (z + \beta)/(1 + \beta z).$$

Well known methods (see for instance op. cit., §10.2) show that $f(z)$ is the unique function analytic and in modulus less than unity within $R: |z| < 1$ which takes the value β for $z=0$ and has the derivative $1-\beta^2$ for the value $z=0$. In the notation of Theorem 1 we set $\beta_{nk}=0$; the extremal properties of $f(z)$ indicate that each of the functions $f_2(z), f_3(z), \dots$ is identical with $f(z)$. Thus we have

$$\limsup_{n \rightarrow \infty} [\max |f(z) - f_n(z)|, \text{ for } |z| \leq r < 1]^{1/n} = 0.$$

But the natural comparison sequence, according to the method of proof of Theorem 1 in somewhat simplified form, is found from the Taylor development of $f(z)$;* we take $r_n(z)$ as the sum of the first n terms of this development:

$$\limsup_{n \rightarrow \infty} [\max |f(z) - r_n(z)|, \text{ for } |z| \leq r < 1]^{1/n} = r/\beta,$$

in contrast to the preceding relation.

3. Extensions of Theorem 1; examples. Merely for the sake of simplicity, we chose in Theorem 1 a region R bounded by a single Jordan curve. The theorem and corollary, together with their proofs, remain valid if R is an arbitrary limited region whose boundary consists of a finite number of mutually disjoint Jordan curves. Likewise the C_2 of Theorem 1 may be replaced by a finite number of mutually disjoint Jordan curves interior to R , no one of which separates any other from the boundary of R or separates any two components of the boundary of R . Under these conditions the locus C_λ also consists of a finite number of mutually disjoint Jordan curves in the region S bounded by C_1 and C_2 , except that for certain values of λ the locus C_λ may have a finite number of multiple points, each shared by a finite number of Jordan curves.

The formal statement of this generalization of Theorem 1 lies immediately at hand, and is left to the reader. A number of special cases of this generalization are worth stating explicitly; in each case we use the notation of Problem A.

* We may equally well choose here the α_{nk} as the $(n-1)$ st roots of A^{n-1} , with $A > 1/\beta$. This choice does not alter the relation involving the functions $r_n(z)$.

(i) Let R be $|z| < 1$, each $\beta_{nk} = 0$, the function $f(z)$ analytic for $|z| < r < 1$ but not for $|z| < r'$, with $r' > r$. Then the situation is analogous to that of Taylor's series; we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} [\max |f(z) - f_n(z)|, \text{ for } |z| \leq r_1 < r]^{1/n} &= r_1/r, \\ (19) \quad \limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_n(z)|, \text{ for } |z| < 1]^{1/n} &= 1/r, \\ \limsup_{n \rightarrow \infty} [\max |f_n(z)|, \text{ for } |z| = r_2 > r]^{1/n} &= r_2/r, \quad r_2 < 1. \end{aligned}$$

(ii) Let R be $|z| < 1$; let each $\beta_{nk} = \beta$, interior to R and independent of n and k ; let the function $f(z)$ be analytic in the region

$$|(z - \beta)/(1 - \bar{\beta}z)| < r < 1$$

but not throughout any region

$$|(z - \beta)/(1 - \bar{\beta}z)| < r' > r.$$

This represents a generalization of (i), and we have obvious equations analogous to (19).

(iii) Let R be $|z| < 1$; let the numbers $\beta_{n1}, \dots, \beta_{nn}$ be the first n numbers of the sequence $\beta_1, \beta_2, \dots, \beta_1, \beta_1, \beta_2, \dots, \beta_1, \beta_1, \beta_2, \dots$, with each β_k interior to R ; let the function $f(z)$ be analytic on the set $|p(z)| < r < 1$ but not throughout any set $|p(z)| < r' > r$, where

$$p(z) = \prod_{j=1}^l \frac{z - \beta_j}{1 - \bar{\beta}_j z}.$$

The point set $|p(z)| < r$ is not necessarily connected. The situation is analogous to that of a certain series of interpolation (op. cit., §9.5). The equations corresponding to (19) are

$$\begin{aligned} \limsup_{n \rightarrow \infty} [\max |f(z) - f_n(z)|, \text{ for } |p(z)| \leq r_1 < r]^{1/n} &= r_1/r, \\ (20) \quad \limsup_{n \rightarrow \infty} [\text{l.u.b. } |f_n(z)|, \text{ for } |p(z)| < 1]^{1/n} &= 1/r, \\ \limsup_{n \rightarrow \infty} [\max |f_n(z)|, \text{ for } |p(z)| = r_2 > r]^{1/n} &= r_2/r, \quad r_2 < 1. \end{aligned}$$

(iv) Let R be $|p(z)| < 1$, where $p(z) = q(z - \beta_1)(z - \beta_2) \cdots (z - \beta_l)$; let the numbers $\beta_{n1}, \beta_{n2}, \dots, \beta_{nn}$ be the first n numbers of the sequence $\beta_1, \beta_2, \dots, \beta_1, \beta_1, \beta_2, \dots, \beta_1, \beta_1, \beta_2, \dots$, with each β_k interior to R ; let the function $f(z)$ be analytic on the set $|p(z)| < r < 1$ but not throughout any set $|p(z)| < r' > r$; the set $|p(z)| < r$ is not necessarily connected. The situation

is analogous to that of the series of interpolation related to the Jacobi series (op. cit., §3.4). Equations (20) are valid also in the present case.

(v) Let R be $|z| < 1$; let the set $\beta_{n1}, \beta_{n2}, \dots, \beta_{nn}$ be the roots of $z^n - b_n = 0$, ($|b_n| \leq b < 1$); let the function $f(z)$ be analytic for $|z| \leq r > b$, $r < 1$, but not analytic throughout $|z| = r'$ with $r' > r$. Then equations (19) are valid provided merely $r_1 \geq b$.

(vi) In the statement of Theorem 1, let C_1 and C_2 be arbitrary (satisfying the conditions imposed), and let $V(x, y)$ denote a function harmonic in S , continuous in the corresponding closed region, taking on the values zero and $\gamma < 0$ on C_1 and C_2 , respectively, where γ is so chosen that the integral of the normal derivative of $V(x, y)$ over an analytic Jordan curve separating C_1 and C_2 is 2π . Let the points β_{nk} be uniformly distributed on C_2 with respect to the function conjugate to $V(x, y)$ in S . Then (op. cit., §8.7) all the conditions of Theorem 1 are fulfilled. This situation is a generalization of (v) if $|b_n| = b$.

Theorem 1 can be extended, as we have indicated, by lessening the restrictions on C_1 and C_2 . Still another extension of Theorem 1 (and of the more general results outlined) is obtained by requiring the limit (1) to hold not at every point exterior to C_2 , but to hold at every point exterior to C_2 except at the points of a set T having no limit point exterior to C_2 , and to hold uniformly on any closed set exterior to C_2 having no point in common with T . The points β_{nk} are no longer required to lie on or interior to C_2 , but must lie in R . No modification need be made in the proofs already given to meet this new hypothesis, except that in the proof of such a relation as (13) where C_σ passes through a point of T , we give the proof first with σ replaced by $\sigma_1 > \sigma$, where C_{σ_1} does not pass through a point of T , and then allow σ_1 to approach σ . With this new requirement on (1), it is not always essential to suppose *all* the points β_{nk} interior to the region R , in which $f(z)$ is assumed defined and analytic; methods for the study of the corresponding sequence $r_n(z)$ are developed in the book already referred to (chap. 11); those methods, together with the present ones, apply directly to the study of Problem A.

We state but a single illustration of the remark just made. Let R be the region $|z| < 1$; let the sequence β_1, β_2, \dots lie interior to $|z| = 1$ and approach zero as its limit, and let us identify $\beta_{n1}, \beta_{n2}, \dots, \beta_{nn}$ with $\beta_1, \beta_2, \dots, \beta_n$; let the function $f(z)$ be analytic for $|z| < r < 1$ but not throughout $|z| < r'$ with $r' > r$. If some of the points β_k lie on or exterior to $|z| = r$, the prescription that $f_n(z)$ shall interpolate to $f(z)$ in those points may be interpreted as requiring that $f_n(z)$ shall interpolate to any function, analytic or not, but not depending on n , in those particular points β_k . The equations (19) are valid.

4. Invariant properties of Theorem 1. Problem A as formulated is invariant under an arbitrary one-to-one conformal transformation. Thus each

of the special situations (i)-(vi) yields, by such a transformation, a new result which the reader can easily express in invariant terms. Theorem 1 itself, especially with regard to condition (1),* has no invariant properties that are obvious, but does have certain relations to invariance, as we shall now proceed to show. The following theorem, previously suggested (op. cit., p. 276) for formulation and proof, is analogous to a theorem already established (op. cit., p. 272, Theorem 20):

THEOREM 2. *Let C' be a Jordan curve of the $w(=u+iv)$ -plane, let the points $w=\beta'_{nk}$ lie on or within C' , and let us suppose*

$$(21) \quad \lim_{n \rightarrow \infty} |(w - \beta'_{n1})(w - \beta'_{n2}) \cdots (w - \beta'_{nn})|^{1/n} = e^{U(u,v)}$$

exterior to C' , uniformly on any closed bounded set exterior to C' . Let a bounded region D' containing C' in its interior be transformed conformally and one-to-one into a bounded region D of the $z(=x+iy)$ -plane by the transformation $w=\phi(z)$, $z=\psi(w)$, with C' transformed into the Jordan curve C and the points β'_{nk} transformed into the points $\beta_{nk}=\psi(\beta'_{nk})$ interior to C . Then the limit

$$(22) \quad \lim_{n \rightarrow \infty} |(z - \beta_{n1}) \cdots (z - \beta_{nn})|^{1/n} = e^{W(x,y)}$$

exists in every finite point exterior to C , uniformly on any closed bounded set exterior to C .

We introduce the notation

$$U_n(x, y) = \frac{1}{n} \sum_{k=1}^n \log |\phi(z) - \phi(\beta_{nk})|, \quad U'_n(x, y) = \frac{1}{n} \sum_{k=1}^n \log |z - \beta_{nk}|, \\ U''_n(x, y) = \frac{1}{n} \sum_{k=1}^n \log \left| \frac{\phi(z) - \phi(\beta_{nk})}{z - \beta_{nk}} \right|,$$

whence $U_n(x, y) = U'_n(x, y) + U''_n(x, y)$. Let Γ denote an arbitrary analytic Jordan curve in D containing C in its interior. Then we have (op. cit., p. 266, Lemma IV) for (x, y) exterior to Γ

$$U'_n(x, y) = \frac{1}{2\pi} \int_{\Gamma} \left(U'_n \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U'_n}{\partial \nu} \right) ds.$$

* Thus if the points β_{nk} are the n roots of unity, equation (1) holds exterior to $C_2: |z|=1$ with $V_1(x, y) = \log |z|$. Under the transformation $z=(w-\beta)/(1-\bar{\beta}w)$, with $|\beta|<1$, the points β_{nk} correspond to the roots of the equation $[(w-\beta)/(1-\bar{\beta}w)]^n - 1 = 0$, and the analogue of (1) is for $|w|>1$

$$\lim_{n \rightarrow \infty} \left| \left[\left(\frac{w-\beta}{1-\bar{\beta}w} \right)^n - 1 \right] \frac{(1-\bar{\beta}w)^n}{1-(-\bar{\beta})^n} \right|^{1/n} = |w-\beta|.$$

The function $U_n''(x, y)$ is harmonic without exception on and interior to Γ (when suitably defined in the points $z = \beta_{nk}$); so we have for (x, y) in D exterior to Γ

$$0 = \frac{1}{2\pi} \int_{\Gamma} \left(U_n'' \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U_n''}{\partial \nu} \right) ds;$$

by addition we write for (x, y) in D exterior to Γ

$$(23) \quad U_n'(x, y) = \frac{1}{2\pi} \int_{\Gamma} \left(U_n \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U_n}{\partial \nu} \right) ds.$$

By hypothesis (21) holds; so the function $U_n(x, y)$ approaches uniformly on Γ the function $U(x, y)$, the transform in the (x, y) -plane of the function $U(u, v)$ in the w -plane; moreover the derivatives of $U_n(x, y)$ on Γ approach uniformly the corresponding derivatives of $U(x, y)$; so by (23) the limit (22) exists, with the relation

$$(24) \quad W(x, y) = \frac{1}{2\pi} \int_{\Gamma} \left(U \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U}{\partial \nu} \right) ds,$$

where it is understood that Γ shall be chosen to contain C but not (x, y) in its interior. Equation (22) is first proved for (x, y) exterior to Γ but interior to D ; however (see op. cit., p. 266) the sequence $U_n'(x, y)$ is a normal family of harmonic functions in the region exterior to C ; when (22) holds in a sub-region, that relation holds uniformly on any closed bounded set exterior to C . Theorem 2 is established.

Theorem 2 extends at once to the more general situation outlined at the beginning of §3.

The significance of Theorem 2 in connection with Theorem 1 lies in two remarks. (i) Although condition (1) is not itself invariant under conformal transformation, and to that extent is unsuited to a discussion of Problem A, condition (1) is shown by Theorem 2 to have certain properties related to invariance, and thereby to be a not unreasonable hypothesis to use. Thus the geometric configuration of Theorem 1 may be subjected to a transformation which carries the closed interior of C_1 into the closed interior of another Jordan curve C_1' conformally and one-to-one. Theorem 1 applies also to the new configuration. (ii) If there is given a region R of simple or multiple connectivity, such that a single Jordan curve or a set of Jordan curves C_2 contains the points β_{nk} not on C_2 in its interior with (1) satisfied, but if R is infinite or if the boundary of R consists not of Jordan curves but of a finite number of other continua, none of which is a single point, then R can be

mapped conformally onto a finite region bounded by a finite number of mutually disjoint analytic Jordan curves, so that condition (1) persists in character, and hence the extension of Theorem 1 applies.

5. Invariant formulation of Theorem 1. Even though Theorem 1 itself is not expressed in form invariant under arbitrary one-to-one conformal transformation, an equivalent result can be so expressed with relative ease, as we shall now proceed to indicate. But our immediate methods apply rather to Theorem 1 itself than to the extension of Theorem 1 to multiply-connected regions R .

THEOREM 3. *Let R be a simply connected region of the extended plane whose boundary C_1 consists of more than two points, and let the function $w = \phi(z)$ map R conformally and one-to-one onto $|w| < 1$. Let C_2 be a Jordan curve interior to R , let C_2 separate the points β_{nk} not lying on C_2 itself from C_1 , and let*

$$(25) \quad \lim_{n \rightarrow \infty} \left| \frac{[\phi(z) - \phi(\beta_{n1})] \cdots [\phi(z) - \phi(\beta_{nn})]}{[\bar{\phi}(\beta_{n1})\phi(z) - 1] \cdots [\bar{\phi}(\beta_{nn})\phi(z) - 1]} \right|^{1/n} = e^{U(x,y)}$$

hold at every point of the annular region S bounded by C_1 and C_2 , uniformly on any closed set interior to S . Let the function $U(x, y)$ be continuous in \bar{S} and take the constant value γ on the curve C_2 . We denote generically by C_λ the locus $U(x, y) = \lambda$, ($\gamma < \lambda < 0$), in R , so that C_λ is a Jordan curve separating C_1 and C_2 ; we denote by R_λ the region bounded by C_λ containing C_2 in its interior, and by \bar{R}_λ the closure of R_λ .

Let the function $f(z)$ be analytic throughout the interior of R_ρ but not throughout the interior of any $R_{\rho'}$, ($\rho' > \rho$). In the notation of Problem A, the sequence $f_n(z)$ converges uniformly to $f(z)$ on any closed set interior to R_ρ . Moreover we have (for $\gamma < \sigma < \rho$) equations (2) and (3).

The functions harmonic in R except in the points β_{nk} ,

$$U_n(x, y) = \frac{1}{n} \log \left| \frac{[\phi(z) - \phi(\beta_{n1})] \cdots [\phi(z) - \phi(\beta_{nn})]}{[\bar{\phi}(\beta_{n1})\phi(z) - 1] \cdots [\bar{\phi}(\beta_{nn})\phi(z) - 1]} \right|,$$

when suitably defined on C_1 are all continuous in the two-dimensional sense on C_1 , except of course that the functions need not be defined exterior to R , and they take the value zero on C_1 . Their uniform convergence on a curve C_λ therefore implies their uniform convergence in the closed region bounded by C_1 and C_λ ; so $U(x, y)$ also is continuous in the two-dimensional sense on C_1 and vanishes there. Of course $U_n(x, y)$ is negative in R , and indeed by the hypothesis on the β_{nk} is uniformly bounded from zero on any closed set interior to S ; so the relation $\gamma < 0$ can be made a matter of proof rather than hypothesis.

Our discussion of Theorem 3 is quite similar to the proof of Theorem 2. Let us transform R conformally without change of notation so that it becomes the interior of an arbitrary Jordan curve C_1 . We introduce the notation

$$U_n'(x, y) = \frac{1}{n} \sum_{k=1}^n \log |z - \beta_{nk}|,$$

$$U_n''(x, y) = \frac{1}{n} \sum_{k=1}^n \log \frac{|\phi(z) - \phi(\beta_{nk})|}{|[z - \beta_{nk}][\bar{\phi}(\beta_{nk})\phi(z) - 1]|},$$

whence $U_n(x, y) = U_n'(x, y) + U_n''(x, y)$.

The function $U_n''(x, y)$, when a suitable definition is provided in the points β_{nk} , is harmonic throughout the interior of R ; so if Γ_2 is an analytic Jordan curve containing C_2 in its interior, but to which (x, y) is exterior, we have (op. cit., p. 265, Lemma III) for (x, y) either in S or on or exterior to C_1

$$0 = \frac{1}{2\pi} \int_{\Gamma_2} \left(U_n'' \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U_n''}{\partial \nu} \right) ds,$$

where ν indicates the interior normal for Γ_2 and the integral is taken in the clockwise sense. Under these circumstances we also have (op. cit., p. 266, Lemma IV) for (x, y) either in S or even on or exterior to C_1

$$U_n'(x, y) = \frac{1}{2\pi} \int_{\Gamma_2} \left(U_n' \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U_n'}{\partial \nu} \right) ds,$$

whence for (x, y) anywhere exterior to C_2 ,

$$(26) \quad U_n'(x, y) = \frac{1}{2\pi} \int_{\Gamma_2} \left(U_n \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U_n}{\partial \nu} \right) ds.$$

The sequence $U_n(x, y)$ converges uniformly to $U(x, y)$ on Γ_2 , and the derivatives of $U_n(x, y)$ converge uniformly on Γ_2 to the corresponding derivatives of $U(x, y)$; so it follows from (26) that $U_n'(x, y)$ converges at every finite point exterior to C_2 , uniformly on any closed limited set exterior to C_2 , to the function

$$(27) \quad U'(x, y) = \frac{1}{2\pi} \int_{\Gamma_2} \left(U \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U}{\partial \nu} \right) ds,$$

where it is understood that Γ_2 is so chosen that (x, y) lies exterior to Γ_2 , and C_2 interior to Γ_2 . With this understanding, the functions $U_n'(x, y)$ and $U'(x, y)$ defined by (26) and (27) are harmonic at every finite point of the plane even exterior to C_1 , and are independent of the particular curve Γ_2 (depending on (x, y)) which is chosen.

Let Γ_1 denote an arbitrary analytic Jordan curve containing in its interior both C_2 and the point (x, y) . Then we have

$$U_n''(x, y) = \frac{1}{2\pi} \int_{\Gamma_1} \left(U_n'' \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U_n''}{\partial \nu} \right) ds,$$

where ν indicates exterior normal for Γ_1 and the integral is taken in the counterclockwise sense. We also have (op. cit., p. 265, Lemma II)

$$0 = \frac{1}{2\pi} \int_{\Gamma_1} \left(U_n' \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U_n'}{\partial \nu} \right) ds,$$

whence for (x, y) interior to Γ_1 ,

$$(28) \quad U_n''(x, y) = \frac{1}{2\pi} \int_{\Gamma_1} \left(U_n \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U_n}{\partial \nu} \right) ds.$$

The sequence $U_n(x, y)$ converges uniformly to $U(x, y)$ on Γ_1 , and the various derivatives of $U_n(x, y)$ converge uniformly on Γ_1 to the corresponding derivatives of $U(x, y)$; so it follows from (28) that $U_n''(x, y)$ converges at every point interior to C_1 , uniformly on any closed set interior to C_1 , even interior to C_2 , to the function

$$(29) \quad U''(x, y) = \frac{1}{2\pi} \int_{\Gamma_1} \left(U \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U}{\partial \nu} \right) ds.$$

It is of course understood that Γ_1 is chosen interior to C_1 , with both (x, y) and C_2 in its interior. With this understanding, the functions $U_n''(x, y)$ and $U''(x, y)$ expressed by (28) and (29) are analytic throughout the interior of C_1 , and are independent of the particular curve Γ_1 (depending on (x, y)) chosen.

The function $U'(x, y)$, harmonic at every point of the plane exterior to C_2 , can now be identified with the function $V_1(x, y)$ of Theorem 1. From

$$U(x, y) = U'(x, y) + U''(x, y),$$

valid interior to S , and from the continuity of $U(x, y)$ and $U'(x, y)$ on C_1 , it follows that $U''(x, y)$ when suitably defined on C_1 also is continuous on C_1 , and takes on the values $-U'(x, y)$ there. Then $U''(x, y)$ is precisely the negative of the function $V_2(x, y)$ of Theorem 1. That is to say, we have shown that under the conditions of Theorem 3 with R the interior of a Jordan curve, the hypothesis of Theorem 1 is satisfied, with $V(x, y)$ of Theorem 1 equal to $U(x, y)$ of Theorem 3; this first yields a proof* of Theorem 3, and second

* A much shorter proof of Theorem 3, which however does not tend to show the equivalence of Theorems 1 and 3, can be given from Theorem 1 by use of the substitution $w = \phi(z)$ in (25).

shows part of the equivalence of Theorems 1 and 2. The complete equivalence of Theorems 1 and 2 will be established by our showing now that the hypothesis of Theorem 1 implies condition (25).

We interpret $U_n''(x, y)$ as the unique function harmonic in R and continuous in the corresponding closed region which equals $-U_n'(x, y)$ on C_1 . By hypothesis* the functions $U_n'(x, y)$ converge uniformly on C_1 to the function $V_1(x, y)$; then the functions $U_n''(x, y)$ converge uniformly on C_1 to the function $-V_1(x, y)$, and hence converge uniformly in the closed region $R+C_1$, to some function $-V_2(x, y)$ harmonic interior to R , continuous in $R+C_1$, and equal to $-V_1(x, y)$ on C_1 . Then the functions $U_n(x, y)$ converge uniformly on any closed set interior to S , to the function $V_1(x, y) - V_2(x, y)$. Consequently, equation (25) is satisfied with $U(x, y)$ equal to the function $V(x, y)$ of Theorem 1, as we desired to show.

Theorem 3, like Theorem 1, applies without further change in proof even if C_2 consists no longer of a single Jordan curve but of several mutually disjoint Jordan curves interior to R , no one of which separates any other from C_1 or separates any two of the components of C_1 ; of course C_2 must separate the β_{nk} not lying on C_2 from C_1 ; the region S is bounded by C_1 and C_2 . The expression of the examples (i)–(vi) in invariant form already suggested is the formulation of several special cases of this extension of Theorem 3.

To Theorem 1 corresponds an expression in form invariant under conformal transformation, namely Theorem 3. Similarly the extension of Theorem 1 to a multiply-connected region R can be expressed in a form invariant under conformal mapping, provided that the connectivity of R is finite and that no component of the boundary C_1 of R consists of a single point; we continue the lighter conditions on C_2 . But here we replace condition (25) by the condition that

$$(30) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n G(x, y; \beta_{nk}) = U(x, y),$$

uniformly on any closed set in the region S bounded by C_1 and C_2 , where $G(x, y; \beta)$ denotes generically Green's function for R with pole in the point β interior to R , and with running coordinates x and y . Condition (30) is a generalization of condition (25), for if R is simply-connected we have the relation

* In the hypothesis of Theorem 1 it is sufficient to assume that (1) holds uniformly merely in S , by virtue of the equation

$$U_n'(x, y) = \frac{1}{2\pi} \int_{\Gamma_2} \left(U_n' \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U_n'}{\partial \nu} \right) ds$$

used in the proof of Theorem 3.

$$G(x, y; \beta_{nk}) = \log \left| \frac{\phi(z) - \phi(\beta_{nk})}{\phi(\beta_{nk})\phi(z) - 1} \right|;$$

the right-hand member is harmonic interior to R except at β_{nk} , is continuous and equal to zero on C_1 , and when diminished by $\log |z - \beta_{nk}|$ is bounded in the neighborhood of the point $z = \beta_{nk}$.

The methods already set forth above show that condition (30) implies the hypothesis of Theorem 1 extended, provided R is a limited region bounded by a finite number of mutually disjoint Jordan curves, and that conversely condition (30) is a consequence of the hypothesis of Theorem 1 extended. We do not emphasize (30) further, however, for it is apparently much more difficult to apply than (25), in the absence of a simple formula for $G(x, y; \beta_{nk})$ when R is multiply-connected.

A consequence of the remark just made is that Theorem 1 extends not merely to a region R bounded by a finite number of mutually disjoint Jordan curves, but also to an arbitrary region R of finite connectivity each component of whose boundary C_1 consists of more than a single point; we still suppose C_2 to consist of a finite number of mutually exterior Jordan curves which separate each of the points β_{nk} not lying on C_2 from the point at infinity. If R is finite, our hypothesis (1) implies, by the reasoning already given in connection with (25) and (30), that equation (30) is valid uniformly on any closed set in S ; consequently Theorem 3 in its extended form applies, and so also does the conclusion of Theorem 1. If R is infinite, we may replace (1) by the condition

$$(31) \quad \lim_{n \rightarrow \infty} \left| \frac{(z - \beta_{n1}) \cdots (z - \beta_{nn})}{(z - \beta)^n} \right|^{1/n} = \frac{e^{V_1(x, y)}}{|z - \beta|},$$

where β is an arbitrary fixed point separated by C_2 from the point at infinity. The function

$$W_n(x, y) = \frac{1}{n} \log \left| \frac{(z - \beta_{n1}) \cdots (z - \beta_{nn})}{(z - \beta)^n} \right|$$

is harmonic even at infinity, when suitably defined there, and the sequence $W_n(x, y)$ converges to the harmonic function $V_1(x, y) - \log |z - \beta|$ (also suitably defined at infinity), uniformly on any closed set *bounded or unbounded* exterior to C_2 . Denote by $g_n(x, y)$ the function harmonic interior to R , continuous in the corresponding closed region, which coincides with $W_n(x, y)$ on C_1 ; the sequence $g_n(x, y)$ converges uniformly on C_1 , and hence converges uniformly in the closed region $R + C_1$, to a function $g(x, y)$ harmonic

in R , continuous in $R+C_1$, equal to $V_1(x, y) - \log |z - \beta|$ on C_1 . We obviously have in the notation of (30)

$$\frac{1}{n} \sum_{k=1}^n G(x, y; \beta_{nk}) = W_n(x, y) - G(x, y; \beta) - g_n(x, y);$$

so equation (30) is satisfied uniformly on any closed set in S with

$$(32) \quad U(x, y) = V_1(x, y) - \log |z - \beta| - G(x, y; \beta) - g(x, y).$$

Consequently Theorem 3 in its extended form applies, and so also does the conclusion of Theorem 1, if we identify $U(x, y)$ as defined by (32) with the function $V(x, y)$ of Theorem 1.

Of course the Corollary to Theorem 1 has an exact analogue in the situation of Theorem 3 extended.

6. Supplementing a given incomplete sequence β_{nk} . It is to be noted that such relations as (2) and (3) involve the superior limit as n takes on *all* the values 1, 2, 3, \dots . Our proofs remain essentially valid if the β_{nk} are defined merely for an infinite sequence of indices n_j , ($j=1, 2, \dots$), with $n_{j+1} > n_j$, provided the difference $n_{j+1} - n_j$ is bounded. But the proofs are no longer valid if the difference $n_{j+1} - n_j$ is not bounded, and (in the absence of specific examples) the analogy with Taylor's series suggests that the conclusions do not remain true. It seems therefore of interest to be able to start with a set β_{nk} satisfying (1) for a suitable sequence of indices n , and to enlarge the set so that (1) is fulfilled for the entire sequence $n=1, 2, \dots$. Methods of solving this problem lie now at hand, as we proceed to indicate.

By our present hypothesis, namely (1) for a suitably chosen sequence of indices n , the function $V_1(x, y)$ is harmonic at every point exterior to C_2 . We define $U_n(x, y)$ by means of (4). Let Γ_2 be an analytic Jordan curve containing C_2 in its interior, but to which (x, y) is exterior. Then we have (op. cit., p. 266, Lemma IV)

$$U_n(x, y) = \frac{1}{2\pi} \int_{\Gamma_2} \left(U_n \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial U_n}{\partial \nu} \right) ds,$$

where the integral is taken in the clockwise sense and ν indicates interior normal for Γ_2 . The function $U_n(x, y)$ approaches $V_1(x, y)$ uniformly on Γ_2 , and the derivatives of $U_n(x, y)$ approach uniformly the corresponding derivatives of $V_1(x, y)$; so we have for (x, y) exterior to Γ_2

$$V_1(x, y) = \frac{1}{2\pi} \int_{\Gamma_2} \left(V_1 \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial V_1}{\partial \nu} \right) ds.$$

By the harmonic character of $V_2(x, y)$ on and within Γ_2 , we may write (op. cit., p. 265, Lemma III) for (x, y) exterior to Γ_2

$$0 = \frac{1}{2\pi} \int_{\Gamma_2} \left(V_2 \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial V_2}{\partial \nu} \right) ds;$$

so for (x, y) exterior to Γ_2 we have

$$V_1(x, y) = \frac{1}{2\pi} \int_{\Gamma_2} \left(V \frac{\partial \log r}{\partial \nu} - \log r \frac{\partial V}{\partial \nu} \right) ds.$$

If C_2 is an analytic Jordan curve, this integral can be taken over C_2 itself; by the constancy of $V(x, y)$, now assumed on C_2 , we have for (x, y) exterior to C_2

$$(33) \quad V_1(x, y) = \frac{-1}{2\pi} \int_{C_2} \log r \frac{\partial V}{\partial \nu} ds.$$

Even if the Jordan curve C_2 is not analytic, equation (33) is valid if the integral is taken in an extended sense (op. cit., §7.6). If the points β'_{nk} are uniformly distributed on C_2 with respect to the parameter σ , where

$$d\sigma = - \frac{\partial V}{\partial \nu} ds,$$

it follows from (33) and the equation

$$\int_{C_2} d\sigma = 2\pi,$$

a consequence of (5), that

$$(34) \quad \lim_{n \rightarrow \infty} | (z - \beta'_{n1}) \cdots (z - \beta'_{nn}) |^{1/n} = e^{V_1(x, y)}$$

for (x, y) exterior to C_2 , uniformly on any closed limited set exterior to C_2 .

If now the given β_{nk} do not appear in (1) for every n , we need merely set $\beta_{nk} = \beta'_{nk}$ for the omitted values of n . Then the new set β_{nk} is defined for every n , and it follows from (1) and (34) that (1) holds uniformly on any closed bounded set exterior to C_2 , when n takes all the values 1, 2, 3, \dots . Such equations as (2) and (3) apply to the new set β_{nk} .

These remarks on supplementing a given incomplete sequence β_{nk} apply without essential change to the more general situation outlined at the beginning of §3.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

MEAN MOTIONS AND ALMOST PERIODIC FUNCTIONS*

BY

PHILIP HARTMAN

Introduction. A continuous function $F(t) = U(t) + iV(t)$, $(-\infty < t < +\infty)$, is said to possess a mean motion μ if it has a representation of the form

$$(1) \quad F(t) = r(t) \exp 2\pi i \phi(t), \quad -\infty < t < +\infty,$$

such that $r(t)$, $\phi(t)$ are real-valued continuous functions and

$$(2) \quad \phi(t)/t \rightarrow \mu (\phi(t) = \mu t + o(t)) \quad t \rightarrow \infty.$$

The problem of the existence and determination of this constant μ for functions $F(t)$ of the type

$$(3) \quad F(t) = \sum_{k=1}^n a_k \exp 2\pi i (\Lambda_k t + \alpha_k),$$

where Λ_k , α_k are real and $a_k > 0$, goes back to Lagrange's approximative treatment of the secular perturbations of the major planets. The earliest result in this direction is that if the amplitudes a_k satisfy Lagrange's relation, that is, if for some j ,

$$(4) \quad a_j > a_1 + \dots + a_{j-1} + a_{j+1} + \dots + a_n,$$

so that

$$(5) \quad |F(t)| > c, \quad -\infty < t < +\infty,$$

for a constant $c > 0$, then the mean motion μ exists and

$$(6) \quad \mu = \Lambda_j.$$

Bohl [1] has proved the existence of μ if $n=3$; Weyl [12] has treated the case $n=4$ when the frequencies $\Lambda_1, \dots, \Lambda_n$ are linearly independent. In the case of a general n , it has been shown (Hartman, van Kampen, and Wintner [5]) that if the numbers a_1, \dots, a_n do not satisfy a relation of the type

$$\sum_{k=1}^n e_k a_k = 0, \quad e_k = \pm 1,$$

and if the frequencies $\Lambda_1, \dots, \Lambda_n$ and the amplitudes a_1, \dots, a_n are fixed,

* Presented to the Society, October 30, 1937; received by the editors April 2, 1938.

then the mean motion μ exists whenever the phases $\alpha_1, \dots, \alpha_n$ do not belong to a certain zero set (which may be empty) in the $(\alpha_1, \dots, \alpha_n)$ -space. Actually, this was stated explicitly only in the case that $\Lambda_1, \dots, \Lambda_n$ are linearly independent, but it is clear from the proof that this restriction is unnecessary. It was also shown that if $(\alpha_1, \dots, \alpha_n)$ does not belong to the exceptional set and if the frequencies are linearly independent, then the mean motion μ possesses an explicit integral representation. More recently, Weyl [13] has shown that if the frequencies are linearly independent, then the exceptional zero set is empty.

It is known* that if $F(t)$ is an arbitrary almost periodic function† satisfying the condition (5), then $\phi(t) = \mu t + \omega(t)$, where $\omega(t)$ is almost periodic. Also (Hartman and Wintner [7]), in this case the mean motion possesses an explicit integral representation.

Let

$$(7) \quad f(s) = f(\sigma + it) = u(\sigma, t) + iv(\sigma, t), \quad \alpha < \sigma < \beta; -\infty < t < +\infty,$$

be an analytic almost periodic function in the strip $\alpha < \sigma < \beta$. In this paper the mean motions of the functions

$$(8) \quad F_\sigma(t) = f(\sigma + it)$$

will be investigated. The method will be that of considering σ as a varying parameter, so that a given $F(t)$ is thought of as embedded into a sheaf of functions (8) depending on σ .

According to Jessen [9], there is associated with every function (7) a Jensen function $\psi(\sigma)$, ($\alpha < \sigma < \beta$), such that $\psi(\sigma)$ is convex and if $\psi(\sigma)$ is differentiable at $\sigma = \alpha'$, $\sigma = \beta'$ for $\alpha < \alpha' < \beta' < \beta$, then the frequency $H(\alpha', \beta')$ of the zeros of $f(s)$ in the strip $\alpha' < \sigma < \beta'$ exists and $2\pi H(\alpha', \beta') = \psi'(\beta') - \psi'(\alpha')$, where $\psi' = d\psi/d\sigma$. In §1, it will be proved that the mean motion $\mu(\sigma)$ exists for every σ , ($\alpha < \sigma < \beta$), at which $\psi'(\sigma)$ exists, and $2\pi\mu(\sigma) = \psi'(\sigma)$. Since $\psi(\sigma)$ is convex, it has a derivative at every point σ with the possible exception of a denumerable set. The connection between the mean motion and the derivative of the Jensen function is established by an adaptation of the methods used by Jessen [9] to prove the existence and the properties of the Jensen function. This connection, when combined with simple examples of limit periodic functions mentioned by Jessen [9], show that on the one hand $\mu(\sigma)$ may exist even though $\psi'(\sigma)$ does not, while on the other hand $\mu(\sigma)$ need not exist for all σ .

A criterion is obtained in §2 for the existence of $\mu(\sigma)$ for all σ in the inter-

* This result was conjectured by Wintner and proved by Bohr [3].

† Throughout this paper, almost periodicity will be meant in the sense of Bohr.

val $\alpha < \sigma < \beta$. The criterion, in the case of an arbitrary function (7), is obtained by a generalization of the methods used by Jessen [10] in the study of zeros of those functions (7) having linearly independent Fourier exponents. However, this general criterion takes a simpler form for a large class of analytic almost periodic functions. An application of this simplified criterion shows that if (7) is a trigonometric polynomial

$$(9) \quad f(s) = \sum_{k=1}^n a_k \exp 2\pi(\Lambda_k s + i\alpha_k),$$

then all of the corresponding functions (8), with the possible exception of a finite set of σ , possess mean motions $\mu(\sigma)$ (§3). The question whether the finite set of exceptional σ is necessarily empty will remain open. It will be shown, however, that if the polynomial (9) has a decomposition

$$(10) \quad f(s) = f_1(s) + f_2(s)$$

into the sum of two polynomials which are not both periodic and whose frequencies are contained in linearly independent moduli, then $\mu(\sigma)$ exists for all σ .

In §4 the smoothness of the function μ of σ is discussed. It is shown, in particular, that in the case of a trigonometric polynomial (9) with linearly independent frequencies, $\mu(\sigma)$ is a regular analytic function at every point σ for which there is no relation of the type

$$\sum_{k=1}^n e_k a_k \cdot \exp 2\pi \Lambda_k \sigma = 0, \quad e_k = \pm 1,$$

while $\mu(\sigma)$ possesses p continuous derivatives for $-\infty < \sigma < +\infty$, if $n \geq 3 + 2p$.

In §5 the methods are extended so as to apply to the Riemann ζ -function for $1/2 < \sigma \leq 1$. As is to be expected, $\mu(\sigma)$ exists for all $\sigma > 1/2$ and $\mu(\sigma) \equiv 0$.

1. Mean motions and the Jensen function. In the sequel, it will be supposed that $f(s) \equiv f(\sigma + it) \not\equiv 0$ is a regular almost periodic function in the strip $\alpha < \sigma < \beta$, where $-\infty \leq \alpha < \beta \leq +\infty$. It will first be shown that every function (8) can be represented in the form (1), where the corresponding function $\phi(t)$ is an analytic function of the real variable t .

LEMMA 1. *For every σ in $\alpha < \sigma < \beta$ there exists a unique function $\phi_\sigma(t)$ satisfying the conditions:*

- (i) $\phi_\sigma(t)$ is a regular analytic function of the real variable t for $-\infty < t < +\infty$.
- (ii) $2\pi\phi_\sigma(t) \equiv \arg F_\sigma(t) \pmod{\pi}$ for $-\infty < t < +\infty$.
- (iii) $0 \leq \phi_\sigma(0) < 1/2$.

This lemma has been proved by Bohl [1] for the case of polynomials (9) when the word "continuous" replaces "regular analytic" in (i). This proof, however, is valid for the case of an arbitrary regular function (7). In order to prove (i) itself, consider (where R is the real part)

$$(11) \quad d\phi_\sigma(t)/dt = (1/2\pi)R\{d \log f(\sigma + it)/ds\},$$

if $F_\sigma(t) = f(\sigma + it) \neq 0$, where $\log f(s)$ is any branch of the logarithm of $f(s)$. Elementary considerations show that the function on the right-hand side of (11) is a continuous, even a regular analytic function for all t , including those t for which $F_\sigma(t) = 0$. From this the analyticity of $d\phi_\sigma(t)/dt$ and, consequently, the analyticity of $\phi_\sigma(t)$ are easily deduced.

Now, according to Jessen [9], the function

$$(12) \quad \psi(\sigma; T) = T^{-1} \int_0^T \log |F_\sigma(t)| dt = T^{-1} \int_0^T \log |f(\sigma + it)| dt,$$

for $\alpha < \sigma < \beta$, $0 < T < \infty$, exists and is a continuous function of σ . The functions (12) tend uniformly to a limit function $\psi(\sigma)$ in every closed subinterval of $\alpha < \sigma < \beta$ as $T \rightarrow \infty$,

$$(13) \quad \psi(\sigma) = \lim_{T \rightarrow \infty} \psi(\sigma; T).$$

The limit function $\psi(\sigma)$, which is called the Jensen function associated with $f(s)$, is convex and has the following property: If $N(\alpha', \beta'; T)$, where $\alpha < \alpha' < \beta' < \beta$, denotes the number of zeros of $f(s)$ in the rectangle $\alpha' < \sigma < \beta'$, $0 < t < T$, and if $\psi(\sigma)$ is differentiable at $\sigma = \alpha'$ and $\sigma = \beta'$, then the limit

$$(14) \quad \lim_{T \rightarrow \infty} N(\alpha', \beta'; T)/T = H(\alpha', \beta')$$

exists and

$$(15) \quad 2\pi H(\alpha', \beta') = \psi'(\beta') - \psi'(\alpha'),$$

where $\psi' = d\psi/d\sigma$. The number $H(\alpha', \beta')$ is called the frequency of the zeros of $f(s)$ in the strip $\alpha' < \sigma < \beta'$. Since a convex function is not differentiable at most on an enumerable set of points, the relation (15) holds with the possible exception of a countable set of α', β' in the interval $\alpha < \sigma < \beta$.

It will be shown that these facts concerning Jensen's function can be transformed into corresponding facts concerning mean motions as follows:

THEOREM I. *If $f(\sigma + it)$ is a regular almost periodic function in the strip $\alpha < \sigma < \beta$, where $-\infty \leq \alpha < \beta \leq +\infty$, then for every σ at which $\psi(\sigma)$ has a derivative, the function*

$$F_\sigma(t) = f(\sigma + it)$$

possesses a mean motion $\mu(\sigma)$ and

$$(16) \quad \mu(\sigma) = (1/2\pi)\psi'(\sigma).$$

Proof.[†] Let $\alpha < \alpha_1 < \alpha' < \beta' < \beta_1 < \beta$, and $F_{\alpha'}(t) \neq 0$, $F_{\beta'}(t) \neq 0$ for $-\infty < t < +\infty$. Since $f(s)$ has only a countable set of zeros in the strip $\alpha < \sigma < \beta$, it is clear that $F_{\sigma}(t) = f(\sigma + it) \neq 0$, $(-\infty < t < +\infty)$, with the possible exception of an enumerable set of σ . Let $t=0$ and $t=T$ be such that

$$(17) \quad |f(\sigma)| > c > 0, \quad |f(\sigma + iT)| > c > 0, \quad \alpha' \leq \sigma \leq \beta'.$$

The almost periodicity of $f(s)$ implies that there exists a number $\tau > 0$ such that in every t -interval $[t^*, t^* + \tau]$, $(-\infty < t^* < +\infty)$, of length τ , there exist values of $t=T$ satisfying (17). Since $f(s) \neq 0$ on the boundary of the rectangle $\alpha' \leq \sigma \leq \beta'$, $0 \leq t \leq T$, one has

$$\begin{aligned} 2\pi N(\alpha', \beta'; T) = & \int_0^T \frac{d \log f(\beta' + it)}{ds} dt - \int_0^T \frac{d \log f(\alpha' + it)}{ds} dt \\ & + i \int_{\alpha'}^{\beta'} \frac{d \log f(\sigma + iT)}{ds} d\sigma - i \int_{\alpha'}^{\beta'} \frac{d \log f(\sigma)}{ds} d\sigma. \end{aligned}$$

It follows that if $\log f(s)$ denotes any fixed branch of the logarithm of $f(s)$ in the neighborhood of the line segments $\sigma = \alpha'$, $(0 \leq t \leq T)$, and $\sigma = \beta'$, $(0 \leq t \leq T)$, then

$$(18) \quad 2\pi N(\alpha', \beta'; T)/T = d\Phi(\beta'; T)/d\sigma - d\Phi(\alpha'; T)/d\sigma + P(\alpha', \beta'; T),$$

where

$$(19) \quad \Phi(\sigma; T) = T^{-1} \int_0^T \log f(\sigma + it) dt,$$

and P denotes a remainder term such that

$$(20) \quad |P(\alpha', \beta'; T)| \leq \frac{2C}{Tc} (\beta' - \alpha'),$$

if C denotes the upper bound of $|f'(s)|$ in $\alpha' \leq \sigma \leq \beta'$.

By the Cauchy-Riemann differential equations and (11), one has in the neighborhood of the lines $\sigma = \alpha'$, $(0 \leq t \leq T)$, and $\sigma = \beta'$, $(0 \leq t \leq T)$,

$$(21) \quad \frac{1}{2\pi} d \log |f(\sigma + it)| / d\sigma = (1/2\pi) R \{ d \log f(\sigma + it) / ds \} = d\phi_{\sigma}(t) / dt.$$

Thus, by (12) and (21)

[†] The first part of this proof is a modification of Jessen's proof of (13), (14), and (15); Jessen [9].

$$(22) \quad \frac{1}{2\pi} T d\psi(\alpha'; T)/d\sigma = \phi_{\alpha'}(T) - \phi_{\alpha'}(0).$$

A similar relation holds if α' is replaced by β' .

The function $\Psi(\sigma; T)$ defined by

$$(23) \quad \Psi(\sigma; T) = \psi(\sigma; T) + \frac{C}{Tc} \sigma^2, \quad \alpha_1 < \sigma < \beta_1,$$

possesses a continuous derivative with respect to σ in a neighborhood of $\sigma = \alpha'$ and $\sigma = \beta'$. Now the real part of the function (19) is $\psi(\sigma; T)$, so that by (18), (20), and (23),

$$(24) \quad 2\pi N(\alpha', \beta'; T)/T = d\Psi(\beta'; T)/d\sigma - d\Psi(\alpha'; T)/d\sigma + \rho(\alpha', \beta'; T),$$

where

$$(25) \quad 0 \geq \rho(\alpha', \beta'; T) \geq -\frac{4C}{Tc} (\beta' - \alpha').$$

It follows from (24) and (25) that

$$d\Psi(\beta'; T)/d\sigma \geq d\Psi(\alpha'; T)/d\sigma.$$

This inequality is clearly valid for all points α', β' in the neighborhood of which $\Psi(\sigma; T)$ has a continuous derivative. Thus, for a fixed T satisfying (17), $\Psi(\sigma; T)$ is convex for $\alpha_1 < \sigma < \beta_1$. In virtue of (13) and (23), one has, uniformly in any closed subinterval of $\alpha_1 < \sigma < \beta_1$,

$$\Psi(\sigma; T) \rightarrow \psi(\sigma), \quad T \rightarrow \infty.$$

Since $\psi(\sigma)$ is convex,

$$(26) \quad \lim_{T \rightarrow \infty} d\Psi(\alpha'; T)/d\sigma = \psi'(\alpha'), \quad \lim_{T \rightarrow \infty} d\Psi(\beta'; T)/d\sigma = \psi'(\beta')$$

if $\psi'(\alpha')$ and $\psi'(\beta')$ exist and if T satisfies the condition (17).

On the other hand (22), (23), (26) imply that if T satisfies (17), then

$$(27) \quad \lim_{T \rightarrow \infty} \phi_{\alpha'}(T)/T = (1/2\pi)\psi'(\alpha'), \quad \lim_{T \rightarrow \infty} \phi_{\beta'}(T)/T = (1/2\pi)\psi'(\beta').$$

In virtue of the remark following (17), to show that (16) holds for $\sigma = \alpha'$ and $\sigma = \beta'$, it is sufficient to prove that there exists a constant M such that

$$(28) \quad |\phi_{\sigma}(t) - \phi_{\sigma}(t')| < M, \quad |t - t'| \leq \tau.$$

Let $F_{\sigma}(t) = U_{\sigma}(t) + iV_{\sigma}(t)$. It is seen from the geometrical relation of $\phi_{\sigma}(t)$ to the curve $x = U_{\sigma}(t)$, $y = V_{\sigma}(t)$ that if t', t'' are any two points in an interval in which $F_{\sigma}(t) \neq 0$, then a necessary condition for

$$|\phi_\sigma(t') - \phi_\sigma(t'')| \geq 1/2$$

is that both $U_\sigma(t)$ and $V_\sigma(t)$ vanish in $t' \leq t \leq t''$. Since there exists† an integer N such that the number of zeros of $f(s)$ in any rectangle $\alpha' \leq \sigma \leq \beta'$, $t^* \leq t \leq t^* + \tau$, for $-\infty < t^* < +\infty$, does not exceed N , the statement (28) follows by an application of the following lemma to the set of functions $z(s) = f(s + it^*)$, $(-\infty < t^* < +\infty)$:

LEMMA 2. *Let Q be an open set containing the closed rectangle $S: \alpha_1 \leq \sigma \leq \beta_1$, $0 \leq t \leq \tau$, and let $\alpha_1 \leq \sigma_0 \leq \beta_1$. For every set Σ of functions $z(s) = x(\sigma, t) + iy(\sigma, t)$ which are regular and uniformly bounded in any closed subset of Q and which do not possess the function $z(s) \equiv 0$ as a limit function, there exists an integer K such that the number of zeros of either $x(\sigma_0, t)$ or $y(\sigma_0, t)$ on the interval $0 \leq t \leq \tau$ does not exceed K .*

This lemma is an immediate consequence of well known properties of normal families.

This completes the proof that if $\psi'(\alpha')$ exists and that if $F_{\alpha'}(t) \neq 0$ for all t , then $\mu(\alpha')$ exists and $2\pi\mu(\alpha') = \psi'(\alpha')$. Actually, the condition $F_{\alpha'}(t) \neq 0$ is not needed. The existence of $\psi'(\alpha')$ implies that $n(\alpha'; T)/T \rightarrow 0$, $T \rightarrow \infty$, where $n(\alpha'; T)$ is the number of zeros of $F_{\alpha'}(t)$ in the interval $0 < t < T$. Suppose that $F_{\alpha'}(t)$ has a zero of order k at $t = t_0$ and that $\eta > 0$ is such that $f(s)$ has no other zeros in $|s - (\alpha' + it_0)| \leq \eta$. Then

$$d \log f(s)/ds = k/[s - (\alpha' + it_0)] + g(s),$$

where $g(s)$ is regular in $|s - (\alpha' + it_0)| \leq \eta$. Thus,‡

$$R\{d \log f(\alpha' + it)/ds\} = R\{g(\alpha' + it)\},$$

so that integration from $\alpha' + i(t_0 - \eta)$ to $\alpha' + i(t_0 + \eta)$ along a semicircle ($\sigma \geq \alpha'$) gives

$$(29) \quad R\left\{-i \int [d \log f(s)/ds] ds\right\} = \phi_{\alpha'}(t_0 + \eta) - \phi_{\alpha'}(t_0 - \eta) + k\pi.$$

Denoting by $L(\alpha'; T)$ the real part of $\{if[d \log f(s)/ds] ds\}$ where the integral extends from $(\alpha' + i0)$ to $(\alpha' + iT)$ along a path consisting of segments of the line $\sigma = \alpha'$ and semicircles ($\sigma \geq \alpha'$) in which there are no zeros other than those on the line $\sigma = \alpha'$, it follows from (15) and the differentiability of $\psi(\sigma)$ at $\sigma = \alpha'$ that

$$L(\alpha'; T)/T \rightarrow \psi'(\alpha'),$$

$$T \rightarrow \infty.$$

† Cf. Bohr and Jessen [4]; or Jessen [9, Lemma 1].

‡ Cf. Lemma 1.

In virtue of (29), (11), and the fact that $n(\alpha'; T)/T \rightarrow 0$, $T \rightarrow \infty$, the mean motion $\mu(\alpha')$ exists and is equal to $\psi'(\alpha')/2\pi$.

2. A criterion for the existence of $\mu(\sigma)$ for every σ . It is clear from the proof of Theorem I that if $N(\alpha', \beta'; T)/T$, $n(\alpha'; T)/T$, $n(\beta'; T)/T$ each has a limit as $T \rightarrow \infty$ and if $\mu(\beta')$ exists, then $\mu(\alpha')$ exists. If, in particular, $N(\alpha', \beta'; T)$ has a limit for every α', β' , $\alpha < \alpha' < \beta' < \beta$, and if $n(\sigma; T)/T \rightarrow 0$, $T \rightarrow \infty$, for every σ , then $\mu(\sigma)$ exists for all σ , ($\alpha < \sigma < \beta$), and the frequency (14) satisfies

$$(30) \quad H(\alpha', \beta') = \mu(\beta') - \mu(\alpha')$$

for every α', β' . In order to investigate under what conditions there are no exceptional α', β' , it is convenient to consider the function $Z(\theta_1, \theta_2, \dots; \sigma)$ defined for every σ , ($\alpha < \sigma < \beta$), on a finite or infinite dimensional θ -torus Θ and associated with $F_\sigma(t)$ in the usual manner (Bohr [2]). Consider those functions $f(s)$ for which there exists a finite or infinite sequence of real, linearly independent numbers $\lambda_1, \lambda_2, \dots$ and, correspondingly, Θ represents a finite or infinite dimensional torus† on which the continuous function $Z(\theta_1, \theta_2, \dots; \sigma)$ is defined for each $F_\sigma(t)$ such that

$$(31) \quad F_\sigma(t) = Z(\lambda_1 t, \lambda_2 t, \dots; \sigma), \quad -\infty < t < +\infty,$$

where the numbers $\lambda_i t$ in (31) are reduced modulo 1. This restriction on $f(s)$ excludes, for example, limit periodic functions for which $\mu(\sigma)$ need not exist‡ for all σ .

Suppose that, for every point $(\theta_1, \theta_2, \dots)$ of the torus Θ , the continuous function $Z(\theta_1, \theta_2, \dots; \sigma)$ is a regular analytic function of the real variable σ . Thus, the torus function Z is still defined if σ is replaced by the complex variable $s = \sigma + it$. Suppose further that the relation

$$(32) \quad Z(\theta_1, \theta_2, \dots; \sigma + it) = Z(\theta_1 + \lambda_1 t, \theta_2 + \lambda_2 t, \dots; \sigma), \\ \alpha < \sigma < \beta; -\infty < t < +\infty,$$

is satisfied.

Let $\nu(\theta_1, \theta_2, \dots; \alpha', \beta')$ denote the number of zeros of $Z(\theta_1, \theta_2, \dots; s)$ in the rectangle

$$(33) \quad S: \alpha' < \sigma < \beta', \quad 0 < t < 1.$$

Then $\nu(\theta_1, \theta_2, \dots; \alpha', \beta')$ is a bounded function on Θ . For otherwise there

† For a theory of limits, measure, and integration on the infinite dimensional torus, see Jessen [10].

‡ Cf. Jessen [9, example 1]. In view of the connection between mean motions and the distribution of zeros, it is easily seen that this function is such that $\mu(0)$ does not exist.

would exist a sequence of points $\{(\theta_1^n, \theta_2^n, \dots)\}$ such that

$$(\theta_1^n, \theta_2^n, \dots) \rightarrow (\theta_1^*, \theta_2^*, \dots), \quad \nu(\theta_1^n, \theta_2^n, \dots; \alpha', \beta') \rightarrow \infty, \quad n \rightarrow \infty.$$

This would imply that $Z(\theta_1^*, \theta_2^*, \dots; s) \equiv 0$, since for reasons of continuity $Z(\theta_1^n, \theta_2^n, \dots; s) \rightarrow Z(\theta_1^*, \theta_2^*, \dots; s)$ holds as $n \rightarrow \infty$ uniformly in S . This is impossible unless $f(s) \equiv 0$. Also, from the continuity of the function Z , it follows by Rouché's theorem that if $Z(\theta_1^0, \theta_2^0, \dots; s)$ has no zeros on the boundary of (33), then $\nu(\theta_1, \theta_2, \dots; \alpha', \beta') = \nu(\theta_1^0, \theta_2^0, \dots; \alpha', \beta')$ for all points in a sufficiently small vicinity of $(\theta_1^0, \theta_2^0, \dots)$ on the torus Θ . Thus, the discontinuity points of ν are among those points $(\theta_1, \theta_2, \dots)$ for which $Z(\theta_1, \theta_2, \dots; s)$ vanishes on the boundary of S .† To insure that the set of discontinuity points of $\nu(\theta_1, \theta_2, \dots; \alpha', \beta')$ is a zero set on Θ assume the following:

(A) The set of all points $(\theta_1, \theta_2, \dots)$ of Θ satisfying either

$$(A, i) \quad Z(\theta_1, \theta_2, \dots; \sigma) = 0$$

for some σ , $\alpha' < \sigma < \beta'$, or at least one of the two relations

$$(A, ii) \quad Z(\theta_1, \theta_2, \dots; \alpha' + it) = 0, \quad Z(\theta_1, \theta_2, \dots; \beta' + it) = 0,$$

for some t , $0 \leq t \leq 1$, is a zero set. (A condition on $Z(\theta_1, \theta_2, \dots; \sigma + 1i)$ similar to (A, i) is unnecessary in virtue of (32).)

Thus, under the condition (A), $\nu(\theta_1, \theta_2, \dots; \alpha', \beta')$ is Riemann integrable over Θ , so that, by the Kronecker-Weyl approximation theorem,

$$(34) \quad \lim_{T \rightarrow \infty} T^{-1} \int_0^T \nu(\lambda_1 t, \lambda_2 t, \dots; \alpha', \beta') dt = \int_{\Theta} \nu(\theta_1, \theta_2, \dots; \alpha', \beta') d\Theta.$$

Since $\nu(\lambda_1 t^*, \lambda_2 t^*, \dots; \alpha', \beta')$ is, by (31), (32), and the definition of ν , the number of zeros of $f(s)$ in the rectangle $\alpha' < \sigma < \beta'$, $t^* < t < t^* + 1$, it is clear from (34) that

$$(35) \quad \lim_{T \rightarrow \infty} N(\alpha', \beta'; T)/T = \int_{\Theta} \nu(\theta_1, \theta_2, \dots; \alpha', \beta') d\Theta.$$

By a slight modification of this argument, it follows from the condition (A, ii) that

$$n(\alpha'; T)/T \rightarrow 0, \quad n(\beta'; T)/T \rightarrow 0 \quad \text{as } T \rightarrow \infty.$$

Thus, if condition (A) is satisfied for all α', β' , $\alpha < \alpha' < \beta' < \beta$, the mean motion $\mu(\sigma)$ exists for all σ and satisfies (30).

† The above arguments are used by Jessen [10, §27], in discussing the zeros of functions $f(s)$ with linearly independent Fourier exponents.

This criterion for the existence of $\mu(\sigma)$ can be transformed into a slightly different form if Θ is a finite dimensional torus. Let Z be a function on a finite dimensional torus Θ , say of dimension $m > 1$. Suppose further that $Z(\theta_1, \theta_2, \dots, \theta_m; \sigma)$ is a regular analytic function of its $m+1$ arguments, in addition to satisfying (32). A necessary and sufficient condition for the conditions (A) to be satisfied for all α', β' is the following:

(B) *There does not exist an $(m-1)$ -dimensional manifold on Θ on which*

$$(36) \quad X(\theta_1, \dots, \theta_m; \sigma) + iY(\theta_1, \dots, \theta_m; \sigma) \equiv Z(\theta_1, \dots, \theta_m; \sigma) = 0$$

for some σ , ($\alpha < \sigma < \beta$).

In order to see this, note that, in virtue of the analyticity of Z in all variables together, the sets involved in (A) are a finite set of manifolds (with possible singularities); so that a necessary and sufficient condition for them to be zero sets is that their dimension numbers be less than m . Now, under the condition (B), the set of points $(\theta_1, \dots, \theta_m; \sigma)$ satisfying (A, i) are manifolds in the $(\theta_1, \dots, \theta_m; \sigma)$ -space with dimension numbers not exceeding $(m-1)$. It follows that the projection of this set on the $(\theta_1, \dots, \theta_m)$ -space Θ is a set of manifolds with dimension numbers not exceeding $(m-1)$, so that it is certainly a zero set. Similar arguments, using (32), show that (B) is necessary as well as sufficient for the condition (A, ii) to be satisfied for all α', β' .

3. Trigonometric polynomials. As an application of the above criterion for the existence of $\mu(\sigma)$ for all σ in an interval, consider a general trigonometric polynomial

$$(37) \quad f(s) = \sum_{k=1}^n a_k \exp 2\pi(\Lambda_k s + i\alpha_k),$$

where Λ_k, α_k are real and $a_k > 0$. It may be supposed that $f(\sigma + it)$ is not a periodic function of t , for this case is trivial. Thus there exist m (greater than 1) linearly independent numbers $\lambda_1, \dots, \lambda_m$ such that

$$(38) \quad \Lambda_k = \sum_{j=1}^m n_{kj} \lambda_j, \quad k = 1, \dots, n,$$

where the n_{kj} are, for $k = 1, \dots, n$ and $j = 1, \dots, m$, integers and the matrix (n_{kj}) is of rank m (less than or equal to n). Thus

$$(39) \quad \begin{aligned} X + iY &\equiv Z = Z(\theta_1, \dots, \theta_m) \\ &= \sum_{k=1}^n a_k \exp 2\pi \left[\Lambda_k \sigma + i \left(\sum_{j=1}^m n_{kj} \theta_j + \alpha_k \right) \right]. \end{aligned}$$

Since $\nu(\theta_1, \theta_2, \dots; \alpha, \beta)$ is uniformly bounded on Θ for any α, β ,

there exists an integer N such that the $N+1$ functions $Z(\theta_1, \dots, \theta_m; \sigma)$, $\partial Z/\partial \sigma, \dots, \partial^N Z/\partial \sigma$ do not vanish simultaneously for $0 \leq \theta_j < 1$, $\alpha \leq \sigma \leq \beta$. If one introduces the jacobian

$$(40) \quad \frac{\partial(X, Y)}{\partial(\sigma, \theta_l)} = 4\pi^2 \sum_{k=1}^n \sum_{p=1}^n a_k a_p \Lambda_k n_{pl} \cdot \exp 2\pi\sigma(\Lambda_k + \Lambda_p) \cdot \cos \left[\sum_{j=1}^m (n_{kj} - n_{pj})\theta_j + (\alpha_k - \alpha_p) \right],$$

(38), (39), and (40) show that

$$(41_0) \quad J_0 \equiv \sum_{l=1}^m \lambda_l \frac{\partial(X, Y)}{\partial(\sigma, \theta_l)} = \left| \frac{\partial Z}{\partial \sigma} \right|^2 = \left(\frac{\partial X}{\partial \sigma} \right)^2 + \left(\frac{\partial Y}{\partial \sigma} \right)^2.$$

Similarly, if one places $X_p = \partial^p X / \partial \sigma^p$, $Y_p = \partial^p Y / \partial \sigma^p$, then

$$(41_p) \quad J_p \equiv \sum_{l=1}^m \lambda_l \frac{\partial(X_p, Y_p)}{\partial(\sigma, \theta_l)} = \left| \frac{\partial^{p+1} Z}{\partial \sigma^{p+1}} \right|^2 = X_{p+1}^2 + Y_{p+1}^2, \quad p = 0, 1, \dots$$

If $\alpha \leq \sigma \leq \beta$, the functions (39), (41₀), \dots , (41 _{$N-1$}) do not vanish simultaneously, so that the set of points $(\theta_1, \dots, \theta_m; \sigma)$, $\alpha \leq \sigma \leq \beta$, at which (39) vanishes is a finite set of disjoint, connected, analytic manifolds, whose dimension numbers do not exceed $(m-1)$. It follows that there are in the interval $\alpha \leq \sigma \leq \beta$ at most a finite number of values σ_0 such that the intersection of these manifolds and the hyperplane $\sigma = \sigma_0$ contains a manifold with a dimension number greater than $(m-2)$. Hence, there is at most a finite number of such exceptional hyperplanes, $-\infty < \sigma < +\infty$, since α, β are arbitrary and the function (39) does not vanish if $|\sigma|$ is sufficiently large; for if the number Λ is chosen so that some of the numbers $\Lambda + \Lambda_1, \dots, \Lambda + \Lambda_n$ are positive and some negative, then $|\exp 2\pi\Lambda\sigma \cdot Z(\theta_1, \dots, \theta_m; \sigma)| \rightarrow \infty$ as $|\sigma| \rightarrow \infty$ uniformly in $(\theta_1, \dots, \theta_m)$.

For arbitrary functions Z of the type (39), this statement is the most general; for example, the hyperplane $\sigma=0$ is exceptional for $Z(\theta_1, \theta_2; \sigma) = \exp 2\pi(\lambda_1\sigma + i\theta_1) + \exp 2\pi(\lambda_2\sigma + i\theta_2)$; also, trigonometric polynomials of the type

$$f(s) = \prod_{j=1}^k (a_{1j} \exp 2\pi\Lambda_{1j}s + a_{2j} \exp 2\pi\Lambda_{2j}s)$$

for properly chosen $a_{1j}, a_{2j}, \Lambda_{1j}, \Lambda_{2j}$ lead to torus functions Z having k exceptional values of σ associated with them.

It follows from the previous section that the mean motion $\mu(\sigma_0)$ exists whenever $\sigma = \sigma_0$ is not an exceptional hyperplane. Thus, the following theorem has been proved:

THEOREM II. *The function*

$$F_{\sigma}(t) = \sum_{k=1}^n a_k \cdot \exp 2\pi\Lambda_k\sigma \cdot \exp 2\pi i(\Lambda_k t + \alpha_k)$$

possesses a mean motion if σ , $(-\infty < \sigma < +\infty)$, does not belong to a certain (possibly empty) finite set.

In some cases, it is certain that the function (37) gives rise to a function Z for which there are no exceptional values of σ . For example, let

$$(42) \quad f(s) = \sum_{k=1}^m a_k \exp 2\pi(\lambda_k s + i\alpha_k), \quad 2 < m < \infty,$$

where again α_k are real, $a_k > 0$, and $\lambda_1, \dots, \lambda_m$ are real linearly independent* numbers. The corresponding torus function is

$$(43) \quad X + iY \equiv Z = Z(\theta_1, \dots, \theta_m; \sigma) = \sum_{k=1}^m a_k \exp 2\pi[\lambda_k \sigma + i(\theta_k + \alpha_k)].$$

It is known† that for any fixed σ , the set of points $(\theta_1, \dots, \theta_m)$ on Θ at which (43) vanishes is either empty or is a finite set of analytic curves if $m=3$; in the case that $m>3$, this set of points, if it is not empty, is an analytic $(m-2)$ -dimensional manifold without singularities or with a finite number of singular curves according as at least one relation of the type

$$(44) \quad \sum_{k=1}^m e_k a_k \exp 2\pi\lambda_k \sigma = 0, \quad e_k = \pm 1,$$

does not or does exist. Thus, by the preceding section, $\mu(\sigma)$ exists for every σ .

It is proved similarly that if the trigonometric polynomial $f(s)$ has the form

$$(45) \quad f(s) = f_1(s) + f_2(s),$$

where $f_1(s), f_2(s)$ are each functions of the type (37), such that not both $f_1(s), f_2(s)$ are periodic and such that the moduli determined by their frequencies are linearly independent, then the corresponding torus function (39) satisfies condition (B) for arbitrary α, β . Thus, one has the following theorem:

THEOREM III. *If the trigonometric polynomial*

$$F(t) = \sum_{k=1}^n a_k \exp 2\pi i(\Lambda_k t + \alpha_k)$$

* For a different proof that the function (42) gives rise to a torus function Z satisfying (A) in the case for $5 \leq m \leq \infty$, cf. Jessen [10, pp. 316-317].

† Hartman, van Kampen, and Wintner [5, pp. 265-266].

can be decomposed into the sum of two trigonometric polynomials F_1, F_2 such that not both F_1, F_2 are periodic and such that the moduli determined by their frequencies are linearly independent, then $F(t)$ possesses a mean motion.

4. **Smoothness of $\mu(\sigma)$.** It is clear from Theorem I that* for any regular analytic almost periodic function $f(s)$, the mean motion $\mu(\sigma)$ is a non-decreasing function on the set on which it is defined. In many cases, certain smoothness properties (for example, differentiability of a given order or analyticity) can be discussed.

Consider first the case (42), where it is known that $\mu(\sigma)$ is defined for all σ . If $5 \leq m \leq \infty$, some of these properties can be deduced from the formula (Jessen [10])

$$(46) \quad H(\alpha, \beta) = \int_{\alpha}^{\beta} G(\sigma; 0, 0) d\sigma,$$

where $G(\sigma; x, y)$ is the density of the asymptotic distribution function of $F_{\sigma}(t)$ with respect to the weight function $|dF_{\sigma}(t)/dt|^2$. By the previous section, (30) holds for all α', β' , so that by (46)

$$(47) \quad \mu(\beta) - \mu(\alpha) = \int_{\alpha}^{\beta} G(\sigma; 0, 0) d\sigma.$$

Now, the function $G(\sigma; x, y)$ is given by the formula†

$$(48) \quad G(\sigma; x, y) = \frac{1}{4\pi^2} \iint \exp i(xu + vy) \cdot X(\sigma; \xi) du dv,$$

where the integral extends over the entire (u, v) -plane, $\xi = u + iv$, and

$$(49) \quad \begin{aligned} X(\sigma; \xi) = & \sum_{k=1}^m |\lambda_k b_k|^2 \prod_{j=1}^m J_0(|b_j \xi|) \\ & + \sum_{\substack{k, l=1 \\ k \neq l}}^m \lambda_k \lambda_l b_k b_l \cdot \prod_{j=1}^m J_0(|b_j \xi|) \cdot J_1(|b_k \xi|) J_1(|b_l \xi|), \end{aligned}$$

where $b_j = b_j(\sigma) = a_j \exp 2\pi \lambda_j \sigma$, the functions J_0, J_1 being the Bessel functions. Using the well known properties of the Bessel functions

$$2dJ_n(w)/dw = J_{n-1}(w) - J_{n+1}(w); |J_n(w)| \leq 1; J_n(w) = O(|w|^{-1/2}), |w| \rightarrow \infty,$$

for $n=0, \pm 1, \dots$, we see that if $m \geq 5+2p$, the function $d^k X/d\sigma^k$ is $O(|\xi|^{-m/2+k})$, so that if $k=1, \dots, p$, then $d^k X/d\sigma^k$ is absolutely integrable

* Compare Jessen [11].

† The formula is obtained by Jessen [10] by methods adapted from Wintner [14].

over the entire $\xi = u + iv$ plane. It follows from (48) that $G(\sigma; x, y)$ has p continuous partial derivatives with respect to σ , so that, by (47), $\mu(\sigma)$ has $p+1$ continuous derivatives in this case. It is clear that if $m = \infty$, then $\mu(\sigma)$ has continuous derivatives of arbitrarily high order.

The formulas (48), (49) were obtained by the use of Fourier transforms, so that it is only possible to decide from them that $G(\sigma; x, y)$ has certain smoothness properties either for all (x, y) or for no (x, y) . On the other hand, it is possible to discuss the existence and smoothness of $G(\sigma; x, y)$ by using methods* recently applied to the density $\delta(\sigma; x, y)$ of the ordinary asymptotic distribution function of $F_\sigma(t)$. These methods apply not only to the case (42) but also to the case of a general trigonometric polynomial (37). It is easily seen that $G(\sigma; x, y)$ can be defined for every point for which $\delta(\sigma; x, y)$ is defined. Also, if for a point $(\sigma; x, y)$ the functions $Z(\theta_1, \dots, \theta_m; \sigma) - (x + iy)$, $\partial(X, Y)/\partial(\theta_k, \theta_j)$, $(k, j = 1, \dots, m)$, do not vanish simultaneously at any point $(\theta_1, \dots, \theta_m)$ of the torus Θ , then $G(\sigma; x, y)$ is defined and is a regular analytic function of its real arguments in a neighborhood of this point. It follows from §2 that if $(\sigma_0; 0, 0)$ is such a point, then $\mu(\sigma)$ exists for all σ sufficiently near to σ_0 , since condition (B) is satisfied if one places $\alpha = \sigma_0 - \epsilon$, $\beta = \sigma_0 + \epsilon$ for a sufficiently small ϵ . On the other hand, it is clear that the considerations of Jessen [10] may be modified to show that (47) holds for $\sigma_0 - \epsilon < \alpha < \beta < \sigma_0 + \epsilon$ (even though the frequencies are not linearly independent). This proves the following:

THEOREM IV. *Let*

$$f(s) = \sum_{k=1}^n a_k \exp 2\pi(\Lambda_k s + i\alpha_k)$$

and let $X + iY \equiv Z = Z(\theta_1, \dots, \theta_m; \sigma)$ be the corresponding function (39). If, for $\sigma = \sigma_0$, the function $Z(\theta_1, \dots, \theta_m; \sigma)$ and the jacobians $\partial(X, Y)/\partial(\theta_j, \theta_k)$, $(j, k = 1, \dots, m)$, do not vanish simultaneously at any point $(\theta_1, \dots, \theta_m)$ of the torus Θ , then $\mu(\sigma)$ exists and is a regular analytic function for all σ sufficiently near to σ_0 .

In the particular case (42), the conditions of Theorem IV are satisfied for all σ for which there is no relation of the type (44). If $m = \infty$ the same is true if (44) is replaced by

$$\sum_{k=1}^n e_k a_k \exp 2\pi\lambda_k \sigma - \sum_{k=n+1}^{\infty} a_k \exp 2\pi\lambda_k \sigma = 0, \quad e_k = \pm 1,$$

* van Kampen and Wintner [8]; Hartman, van Kampen, and Wintner [6]. The formulation of the results in the latter paper can be extended at once to the density $G(\sigma; x, y)$ of the weighted distribution function.

for some n . Summarizing the results for the functions (42), one has the following:

THEOREM V. *If the numbers $\lambda_1, \lambda_2, \dots$ are linearly independent, the function*

$$F_\sigma(t) = \sum_{k=1}^m a_k \exp 2\pi(\lambda_k s + i\alpha_k), \quad 1 \leq m \leq \infty,$$

possesses a mean motion $\mu(\sigma)$. If $3+2p \leq m \leq \infty$, $\mu(\sigma)$ possesses p continuous partial derivatives. If $m < \infty$, $\mu(\sigma)$ is a regular analytic function at every point, with the possible exception of those σ for which there is a relation of the type

$$\sum_{k=1}^m e_k a_k \exp 2\pi\lambda_k \sigma = 0, \quad e_k = \pm 1.$$

Finally, if $m = \infty$, $\mu(\sigma)$ is regular analytic at every point, with the possible exception of those σ for which there is a relation of the type

$$\sum_{k=1}^n e_k a_k \exp 2\pi\lambda_k \sigma - \sum_{k=n+1}^{\infty} a_k \exp 2\pi\lambda_k \sigma = 0, \quad e_k = \pm 1,$$

for some n .

5. Mean motions of the Riemann ζ -function. It is clear from Theorem I, and the fact that $\zeta(s)$ is almost periodic (in the sense of Bohr) and does not vanish for $\sigma > 1$, that the mean motion $\mu(\sigma)$ exists for every $\sigma > 1$. Furthermore, $\mu(\sigma)$ is independent of σ ; since the real part of $\zeta(s)$ does not vanish when σ is sufficiently large, $\mu(\sigma) = 0$ for all $\sigma > 1$.

Although $\zeta(s)$ is not almost periodic in the sense of Bohr for $1/2 < \sigma \leq 1$, the methods developed in §1 can be adapted for this case by using the well known fact that $N(\alpha, \beta; T)/T \rightarrow 0$ as $T \rightarrow \infty$, where $1/2 < \alpha < \beta \leq \infty$ and $N(\alpha, \beta; T)$ denotes the number of zeros of $\zeta(s)$ in the rectangle $\alpha < \sigma < \beta$, $1 < t < T$. Thus, it may be concluded that $\mu(\sigma) \equiv 0$ for $\sigma > 1/2$.

BIBLIOGRAPHY

1. P. Bohl, *Über ein in der Theorie der säkularen Strömungen vorkommendes Problem*, Journal für die reine und angewandte Mathematik, vol. 135 (1909), pp. 189-283.
2. H. Bohr, *Zur Theorie der fastperiodischen Funktionen*, II, Acta Mathematica, vol. 46 (1925), pp. 101-214.
3. ———, *Kleinere Beiträge zur Theorie der fastperiodischen Funktionen*, I, Danske Videnskabskabernes Selskab, Matematisk-Fysiske Meddelelser, vol. 10 (1930), no. 10.
4. H. Bohr and B. Jessen, *Über die Werteverteilung der Riemannschen Zetafunktion*, I, Acta Mathematica, vol. 54 (1930), pp. 1-35.
5. P. Hartman, E. R. van Kampen, and A. Wintner, *Mean motions and distribution functions*, American Journal of Mathematics, vol. 59 (1937), pp. 261-269.

6. ———, *On the distribution functions of almost periodic functions*, American Journal of Mathematics, vol. 60 (1938), pp. 491–500.
7. P. Hartman and A. Wintner, *On the secular constants of almost periodic functions*, Travaux de l'Institut Mathématique de Tblissi, vol. 2 (1937), pp. 37–40.
8. E. R. van Kampen and A. Wintner, *Convolutions of distributions on convex curves and the Riemann zeta-function*, American Journal of Mathematics, vol. 59 (1937), pp. 175–204.
9. B. Jessen, *Über die Nullstellen einer analytischen fastperiodischen Funktion; eine Verallgemeinerung der Jensenschen Formel*, Mathematische Annalen, vol. 108 (1933), pp. 485–516.
10. ———, *The theory of integration in a space of an infinite number of dimensions*, Acta Mathematica, vol. 63 (1934), pp. 249–323.
11. ———, *Om Sekularkonstanter for en næstenperiodisk Funktion*, Matematisk Tidsskrift, B, 1937, pp. 45–48.
12. H. Weyl, *Sur une application de la théorie des nombres à la mécanique statistique et la théorie des perturbations*, Enseignement Mathématique, vol. 16 (1914), pp. 455–467.
13. ———, *Mean motion*, American Journal of Mathematics, vol. 60 (1938), pp. 889–896.
14. A. Wintner, *Upon a statistical method in the theory of diophantine approximations*, American Journal of Mathematics, vol. 55 (1933), pp. 309–331.

THE JOHNS HOPKINS UNIVERSITY,
BALTIMORE, MD.

MAXIMAL ORDERS IN RATIONAL CYCLIC ALGEBRAS OF COMPOSITE DEGREE*

BY
SAM PERLIS

Introduction. A maximal order M of a normal division algebra D over the rational number field may be imbedded[†] in a simple fashion in a maximal order of any normal simple algebra similar to D . When the normal simple algebra has degree greater than two, its class number is unity,[‡] and it can then be shown that all maximal orders of the algebra are obtainable from any one by an inner automorphism of the algebra. Thus it is sufficient to determine a single M of each D in order to determine all maximal orders of all normal simple algebras of degree greater than two over the rational number field. This determination was made by Hull[§] for the case in which the degree n of D is any odd prime, using methods similar to those of Albert^{||} for the case $n=2$. The methods and results of Hull are extended here to the case in which $n=\pi^e$ where π is any odd prime, and also to the case $n=2^e > 2$ provided that D has odd discriminant and has the real number field as splitting field.

More specifically, it will be shown with the aid of the class field theory that each algebra D considered has a suitably normalized cyclic generation, and a maximal order of D will be expressed in terms of a finite number of quantities related to this generation. There are two chief points of difference between the present case and that of prime degree. The quantity σ in the normalized generation (Z, S, σ) is no longer the product of the primes ramified in D , but the product of certain powers of these primes. The exponents on these powers reduce to unity in the case of prime degree. The explicit basis given for the maximal order is similar to that for prime degree

* Presented to the Society, April 9, 1938; received by the editors October 13, 1938. While preparing this paper the author had the privilege of discussing some of its details with Professor A. A. Albert, and is grateful for this guidance and stimulus.

† For the concepts and results on the arithmetic of algebras see M. Deuring, *Algebren, Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 4, no. 1.

‡ M. Eichler, *Bestimmung der Idealklassenzahl in gewissen normalen einfachen Algebren*, *Journal für die reine und angewandte Mathematik*, vol. 176 (1936), pp. 192-202.

§ Ralph Hull, *Maximal orders in rational cyclic algebras of odd prime degree*, these Transactions, vol. 38 (1935), pp. 514-530. For the case $n=2$ see Hull's paper in the same journal, vol. 40 (1936), pp. 1-11. Reference to the first of these papers will be made by the letter H.

|| A. A. Albert, *Integral domains of rational generalized quaternion algebras*, *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 164-176.

except for the appearance of certain rational integral denominators which, again, reduce to unity in the case $n=\pi$.

For algebras of arbitrary degree, the determination may be reduced to the prime-power case if one can express maximal orders in a direct product of two normal division algebras of relatively prime degrees in terms of maximal orders in the two factors. A partial discussion of this direct product theory is given in the final section. The product of two orders, one in each factor, is an order in the direct product of the algebras, and it is shown that this order is maximal if and only if the discriminants of the two algebras are relatively prime. This result holds if any normal simple algebras are used instead of normal division algebras.

1. Cyclic generations and related concepts. A normal division algebra D of degree n over R is a cyclic algebra

$$(1) \quad D = (Z, S, \gamma)$$

and has a basis

$$(2) \quad u^{i-1}z_j, \quad i, j = 1, \dots, n,$$

where (z_1, \dots, z_n) is a basis of the cyclic field Z over R with generating automorphism S , and

$$(3) \quad u^n = \gamma, \quad zu = uz^S$$

for every z of Z .

Since $(Z, S, \gamma) = (Z, S, \gamma\rho^n)$ for any rational number $\rho \neq 0$, it follows that the quantity γ of R may be assumed with no loss of generality to be a rational integer. If we choose for the basis (z_1, \dots, z_n) a minimal basis of Z , then the set of all linear combinations of the n^2 quantities (2) with coefficients rational integers is an order of D . This order is uniquely determined by the cyclic generation (1) of D and is called *the order I in D associated with this generation*. Every order of D , in particular an order I , is contained in a maximal order* of D . We shall obtain an infinite number of normalized cyclic generations of D and for each of the corresponding orders I we shall obtain n distinct maximal orders containing I .

A complete set of invariants of D under change of cyclic generation has been obtained by Hasse† in terms of the norm residue symbol

$$(4) \quad (\gamma, Z | q) = \left(\frac{\gamma, Z}{q} \right) = S^{\nu_q}$$

* Deuring, op. cit., p. 70.

† H. Hasse, *Theory of cyclic algebras over an algebraic number field*, these Transactions, vol. 34 (1932), pp. 171-214.

which is defined for every prime spot q of R . We shall adopt the convention that the integer ν_q is one of $0, 1, \dots, n-1$. For any cyclic algebra $D_1 = (Z_1, S_1, \gamma_1)$ of degree n over R , we have $(\gamma_1, Z_1|q) = S_1^{\nu_q}$, and Hasse has shown that D_1 is equivalent to D if and only if $\nu_1 = \nu_q$ for every q . Hence the ν_q and the degree n form a complete set of invariants of D .

It is known that the norm residue symbol is the identity automorphism, that is, $\nu_q = 0$, for all but a finite number of prime spots $q = q_1, \dots, q_s$. These are precisely the prime spots for which the q -adic extension $D_q = D \times R_q$ is not total matric, and also are characterized as the prime factors of the discriminant of D . These prime spots q_1, \dots, q_s are called the *ramification spots* of D , and a cyclic algebra has at least two ramification spots unless it is total matric. The invariants ν_q satisfy the relations

$$(5) \quad \sum_q \nu_q \equiv 0 \pmod{n}, \quad 2\nu_{q_\infty} \equiv 0 \pmod{n}$$

where q_∞ is the infinite prime spot of R , and these are the only relations between the invariants of an arbitrary cyclic algebra over R . However, a necessary and sufficient condition that a cyclic algebra D of prime-power degree $n = \pi^e$ over R be a division algebra is that at least one of its q -adic extensions be a division algebra, and this is equivalent to the condition that the corresponding invariant ν_q be prime to n . Both of these equivalent conditions follow readily from theorems* that (1) the q -index of $D = (Z, S, \gamma)$ over R is the order of the automorphism S^{ν_q} and thus is

$$(6) \quad n_q = n / (n, \nu_q);$$

and (2) the index of the cyclic algebra D is the least common multiple of all of its q -indices n_q .

Until §5 it will always be assumed that the normal division algebra D has prime-power degree $n = \pi^e > 2$ over R so that there exists a ν_q which is prime to n . From (5) one obtains $\nu_{q_\infty} = 0$ if n is odd (so that in this case q_∞ cannot be a ramification spot q_i), and $\nu_{q_\infty} \equiv 0 \pmod{2}$ if $n = 2^e > 2$. In any case, ν_{q_∞} is not prime to n . Conditions (5) also imply that $s \geq 2$ and that there must be at least two prime spots for which the corresponding invariants are prime to n . Hence we may hereafter let q_1 designate a ramification spot such that $(\nu_{q_1}, n) = 1$ and $q_1 \neq \pi$.

2. Normalized cyclic generations. Three lemmas will now be obtained for use in the proof of Theorem 1 which provides cyclic generations of an especially simple type for the algebra D . The first lemma defines a collection of fields from which the cyclic generation fields of D will be selected.

* Hasse, *ibid.*, Theorem 5, p. 179, and (17.7), p. 203.

LEMMA 1. For any prime $p \equiv 1 \pmod{2n}$ let H_p be the ideal group in R consisting of all principal ideals (r) where r is a rational number prime to p and is an n -ic residue modulo p . Then the class field Z_p corresponding to H_p is cyclic of degree n over R and has conductor p .

If we let G_p be the group of all (r) with r prime to p and let g be a primitive root of p , we shall verify the decomposition

$$(7) \quad G_p = H_p + H_p g + \cdots + H_p g^{n-1}.$$

When $p \equiv 1 \pmod{2n}$, a quantity $\pm g^i$ is an n -ic residue modulo p if and only if i is a multiple of n , whence it follows that the cosets $H_p g^i$ are distinct. For any (r) in G_p we have $r = ab^{-1}$, a and b integers prime to p , $a = g^e + xp$, $b = g^f + yp$ with integers x and y . Then

$$r = \frac{g^e + x_1 p}{g^f + y_1 p} g^{e-f} = r_1 g^{e-f} = \frac{g^e + x_1 p}{g^e + y_1 p} g^{e-f}.$$

If $f \geq e$, the number x_1 is an integer, and we have $r_1 \equiv 1 \pmod{p}$. Otherwise y_1 is integral and again we have the same congruence, so that (r_1) is in H_p and (r) is in $H_p g^{e-f}$, which is one of the cosets displayed. This verifies the decomposition above. The prime p is a generating modulus of the ideal group H_p so that the conductor of H_p , which is the g.c.d. of all the generating moduli, is either p or 1. Then clearly the conductor of H_p , and hence of Z_p , is p ; and since G_p/H_p is cyclic of order n , the field Z_p is cyclic of degree n over R .

Since the next two lemmas depend on the notations of Theorem 1, the latter result will be stated now but not proved until the lemmas have been obtained.

THEOREM 1. Let D be a normal division algebra of prime-power degree $n = \pi^e$ over the rational number field R , and let q_1, \dots, q_s be the finite ramification spots of D and n_i the q_i -index of D , ($i = 1, \dots, s$). Then, if π is odd, there are infinitely many cyclic fields Z of degree n over R such that

- (a) $D = (Z, S, \sigma)$, $\sigma = \prod_{i=1}^s q_i^{n/n_i}$;
- (b) Z has conductor a prime p such that $p \equiv 1 \pmod{n}$, $(p, \sigma) = 1$;
- (c) q_1, \dots, q_s generate prime ideals (q_i) in Z ;
- (d) σ is an n -ic residue modulo p .

If $n = 2^e > 2$ the same results hold provided that D is unramified at the prime spot 2 and at the infinite prime spot q_∞ .

Let v_∞ and v_1, \dots, v_s be the invariants corresponding to q_∞ and the q_i . As we have already seen, our hypotheses imply that $v_\infty = 0$. By (6) we have $n_i = n/(n, v_i)$, and the congruences

$$(8) \quad -v_1 n n_i^{-1} x_i \equiv v_i \pmod{n}, \quad i = 2, \dots, s,$$

have solutions x_i since $(n, v_1) = 1$ and $(n, v_1 n n_i^{-1}) = (n, n n_i^{-1}) = n n_i^{-1} = (n, v_i)$. Note that the x_i are prime to n . Let ζ be a primitive n th root of unity; let

$$(9) \quad \alpha_i = (q_1^{x_i} q_i)^{n/n_i}, \quad i = 2, \dots, s,$$

$$(10) \quad F = R(\zeta), \quad K = F(\alpha_2^{1/n}, \dots, \alpha_s^{1/n}).$$

LEMMA 2. *The field $K_1 = K(q_1^{1/\pi})$ has degree π over K .*

Consider an equation

$$(11) \quad q_1^{c_1} (q_1^{x_2} q_2)^{c_2} \cdots (q_1^{x_s} q_s)^{c_s} = a^n, \quad a \text{ in } F,$$

where the c_i are integers to be determined, and suppose that π is not one of the q_i . Then all the q_i are unramified in F since the discriminant of F is a power* of π . Hence the prime ideal factorization of the quantities in (11) shows that $c_1 + c_2 x_2 + \cdots + c_s x_s$ and c_2, \dots, c_s are all divisible by n , and therefore c_1 is divisible by n . Thus (11) holds only when the exponents c_i are all multiples of n , a property which implies† that the composite of the fields $F(q_1^{1/n})$ and $F([q_1^{x_i} q_i]^{1/n})$ for $i = 2, \dots, s$ is their direct product. These s fields have subfields $F(q_1^{1/\pi})$ and $F(\alpha_i^{1/n})$ for $i = 2, \dots, s$, respectively, and the composite of these subfields must be their direct product. Then the degree of K_1 over K is the degree of $F(q_1^{1/\pi})$ over F , and this is either‡ π or 1. If the degree were 1, then F would contain $q_1^{1/\pi}$, q_1 would be the π th power of an ideal in F , whereas q_1 is prime to π and hence unramified in F . We have proved the lemma for the case in which π is not one of the q_i .

In case π is one of the q_i , we have assumed $\pi > 2$ and may take $\pi = q_2$. Consider an equation of the form (11) with the factor $q_1^{c_1}$ deleted, and obtain $c_2 x_2 + \cdots + c_s x_s \equiv c_3 \equiv \cdots \equiv c_s \equiv 0 \pmod{n}$ since q_1 and q_3, \dots, q_s are unramified in F . Thus $c_2 x_2$ is divisible by n , x_2 is prime to n , and $c_2 \equiv 0 \pmod{n}$. As in the previous paragraph, the composite K_0 of the fields $F([q_1^{x_i} q_i]^{1/n})$ for $i = 2, \dots, s$ is then their direct product and (by Bericht II, p. 43) any cyclic subfield of K_0 has the form

$$(12) \quad F([(q_1^{x_2} q_2)^{d_2} \cdots (q_1^{x_s} q_s)^{d_s}]^{1/n}), \quad d_i \text{ integers.}$$

If $F(q_1^{1/\pi})$ is contained in K_0 , it must have the form (12) so that§

* R. Fricke, *Lehrbuch der Algebra*, 1928, vol. 3, p. 195.

† H. Hasse, *Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper*, Teil II, Jahresbericht der deutschen Mathematiker-Vereinigung, supplementary vol. 6 (1930), p. 43. Parts I and Ia of this article appeared in the Jahresbericht, vols. 35 and 36. These papers will be designated here as Bericht I, Ia, and II.

‡ Bericht II, p. 42, Theorem I.

§ Bericht II, p. 42, Theorem II.

$$q_1^{n/\pi} = c^n q_1^{(x_2 d_2 + \dots + x_s d_s) x \pi d_2 x} \cdot q_2^{d_2 x} \cdot \dots \cdot q_s^{d_s x}$$

with c in F . By considering prime ideal factorizations of the quantities in this equation, we find that $d_2 x, \dots, d_s x$ are divisible by n , $x_2 d_2 x + \dots + x_s d_s x \equiv x_2 d_2 x \equiv n/\pi \pmod{n}$, x_2 is prime to n , $d_2 x \equiv x_0 n/\pi \pmod{n}$. The equation above then takes the form

$$q_1^{n/\pi} = c_0^n q_1^{n/\pi} \pi^{x_0 n/\pi}, \quad c_0^{-n} = \pi^{x_0 n/\pi}.$$

Since x_0 is prime to n we easily obtain $\pi^{n/\pi} = c_0^n$ with c_0 in F and thus have $\pi^{1/\pi}$ in F . Then F must contain the non-normal subfield $R(\pi^{1/\pi})$ whereas F is cyclic and all of its subfields are normal over R . We have shown that $F(q_1^{1/\pi})$ is not contained in K_0 . Then it is not contained in the subfield K of K_0 and the lemma is proved.†

LEMMA 3. *There are infinitely many rational primes p such that $p \equiv 1 \pmod{2n}$, $(p, \sigma) = 1$, and*

- (e) $\alpha_2, \dots, \alpha_s$ are n -ic residues modulo p ;
- (f) q_1^t is an n -ic non-residue modulo p for $t=1, \dots, n-1$.

The field K_1 of Lemma 2 is cyclic of degree greater than 1 over K and is class field to an ideal group H_1 in K . In any ideal class different from the identity class H_1 , we may select an infinite number of prime ideals P which are of degree one, prime to σ , and prime to the different of K over R . An infinite number of rational primes $p = N_{K/R}(P)$ is thus defined. Every such p is prime to σ ; and since the prime ideal factors of p in F must have degree one, it follows that $p \equiv 1 \pmod{n}$. Then $p \equiv 1 \pmod{2n}$ if n is odd.

When $n = 2^e$ we shall make the following additional restrictions in the choice of the ideals P . Let F_2 be the root field over R of the equation $x^{2^n} = 1$ so that F_2 has degree two over F . The field K cannot contain F_2 since then F_2 would have the form (12) which leads to a contradiction. Hence the composite (K, F_2) has degree two over K and is the class field corresponding to an ideal group H_2 in K . We wish to choose ideals P lying outside of H_1 as before but also lying in H_2 . Let these ideal groups have a common generating modulus. Then H_1 and H_2 are collections of ray classes, and we must verify that the ray classes comprising H_2 do not all lie among those comprising H_1 . This fact is clearly true since otherwise $(K, F_2) = K$, $K(\zeta^{1/2}) = K(q_1^{1/2})$, which is impossible. Thus there is a ray class C in H_2 but not in H_1 , and C contains infinitely many prime ideals with the properties of the previous paragraph.

† Since $(q_1^{x_2 \pi})^{1/n_2}$ is in K , this field contains $q_1^{1/\pi}$ if and only if it contains $\pi^{1/\pi}$. Then we see that Lemma 2 is false without the hypothesis that $\pi \neq 2$ when π is one of the q_i . For, if $\pi = 2$, take $n \geq 8$ and see that F , and hence K , contains a primitive eighth root ζ_8 of unity and thus contains $\zeta_8 - \zeta_8^3 = 2^{1/2}$, so that $q_1^{1/2}$ is in K and the lemma fails.

The norms of these ideals are rational primes p such that $p \equiv 1 \pmod{2n}$ since they are unramified in F_2 and their prime ideal factors in F_2 have degree one.

The proofs of properties (e) and (f) are similar to corresponding proofs in H and will be omitted here.* To prove Theorem 1, let p be any prime of Lemma 3 and let Z be the corresponding field Z_p of Lemma 1. Then property (b) of the theorem holds. Property (f) of the last lemma is equivalent to the statement that q_1 is a prime ideal in Z , and property (e) implies that the α_i are in the ideal group H_p corresponding to Z . Expressed in terms of Artin symbols these facts yield

$$(13) \quad (Z/\alpha_i) = I, \quad (Z/q_1)^{x_i n/n_i} = (Z/q_1)^{-n/n_i}.$$

Since x_i is prime to n and the automorphism (Z/q_1) has order n , it follows that $(Z/q_1)^{x_i n/n_i}$ has order n_i . A simple computation shows that (Z/q_1) has order n , which is equivalent to (c). Applying (e) together with (8) and (5), we are led to (d).

The Artin symbol $A = (Z/q_1)$ is a generating automorphism of Z over R , and the equation $S^n = A^{-1}$ defines another generating automorphism S . Then (Z, S, σ) is a cyclic algebra of degree n . A computation following the pattern in H shows that D and (Z, S, σ) have the same invariants, yielding (a) and completing the proof of the theorem.

3. Some properties of Z . Since Z is cyclic over R with conductor p , it is a subfield† of the cyclotomic field $R(\xi)$, where ξ is a primitive p th root of unity. The field $R(\xi)$ is cyclic over R so that Z is its unique subfield of degree n , and Z is thus uniquely determined by its degree n , its prime conductor p , and the property of being an abelian field over R . Write $p = 1 + hn$, and let g be a primitive root of p . Then a normal basis of Z is given by‡

$$\eta_0, \eta_1, \dots, \eta_{n-1}$$

with

$$(14) \quad \eta_i = \xi_i + \xi_{i+n} + \dots + \xi_{i+(h-1)n}, \quad \xi_k = \xi^{g^k},$$

* We may observe that Lemma 3 is actually false without the assumption $\pi \neq 2$ when π is one of the q_i . For, without this assumption we may have $n = 2^s$, $K \geq F(q_1^{1/2}) = K_0$, and $H \leq H_0$, where H and H_0 , respectively, are the ideal groups in F corresponding to the class fields K and K_0 over F . The condition $p \equiv 1 \pmod{n}$ implies that any prime factor P of p in F has degree 1, and condition (e) implies that P is in H and hence in H_0 . Then any prime factor P_0 of P in K_0 has degree 1; hence the quantity $q_1^{1/2}$ of K_0 satisfies $q_1^{1/2} \equiv y \pmod{P_0}$ with y in R . Then $q_1^{n/2} \equiv y^n \pmod{P_0}$ so that we have $q_1^{n/2} \equiv y^n \pmod{p}$, a contradiction with (f).

The falsity of Lemma 3 can be seen to imply the falsity of the conclusions in Theorem 1. Thus the restrictive assumption in Theorem 1 is necessary.

† Bericht I, p. 39.

‡ B. L. van der Waerden, *Moderne Algebra*, 1930, vol. 1, pp. 160 ff.

for $i=0, \dots, n-1$ and $k=0, 1, \dots, p-2$. Hence $Z=R(\eta_i)$ for any i , and a generating automorphism of Z over R is induced by

$$U: \quad \xi \longmapsto \xi^\sigma.$$

Clearly, U is a generating automorphism of the cyclic group $[U]$ of $R(\xi)$ over R , and $[U^n]$ is the group of $R(\xi)$ over Z .

The factorization of p in Z may now be obtained. Define

$$(15) \quad \beta = \prod_{t=0}^{h-1} (1 - \xi^{\sigma^{nt}}).$$

Then β is unaltered by U^n and hence is in Z , and a direct computation shows that $N_{Z|R}(\beta) = p$. The principal ideal $P = (\beta)$ is thus a prime ideal of Z and is a factor of p . But p is completely ramified in the cyclotomic field $R(\xi)$ and hence in the subfield Z , so that $p = P^n$. This fact and Theorem (14) of §8, Bericht Ia, may be used to show that the discriminant of Z over R is p^{n-1} . We thus have

THEOREM 2. *Each field Z of Theorem 1 (and Z_p of Lemma 1) has discriminant p^{n-1} . The factorization of p in Z is*

$$p = P^n, \quad P = (\beta), \quad N_{Z|R}(\beta) = p,$$

where β is given by (15).

The quantity β will be used in the next section when basal elements of maximal orders are defined.

4. Maximal orders in D . The algebra D has the form

$$D = Z + uZ + \dots + u^{n-1}Z, \quad u^n = \sigma,$$

and this generation of D is associated with an order

$$(16) \quad I = Z_0 + uZ_0 + \dots + u^{n-1}Z_0$$

where Z_0 is the maximal order of Z . We shall display n distinct maximal orders in D which contain I . These n orders are defined in terms of n rational integers λ given in

LEMMA 4. *The simultaneous congruences*

$$(17) \quad \lambda^n \equiv \sigma \pmod{p}, \quad \lambda \equiv 0 \pmod{\sigma}$$

have exactly n solutions λ which are incongruent modulo p .

Any solution of the second congruence has the form $\lambda_0\sigma$. If this is substituted in the first congruence, there results

$$(18) \quad \lambda_0^n \equiv \sigma\sigma^n \pmod{p}$$

with $\sigma\sigma_1 \equiv 1 \pmod{p}$. There exists a solution of (18) if and only if* we have $(\sigma\sigma_1^n)^{(p-1)/g} \equiv 1 \pmod{p}$ where $g = (p-1, n)$; then the exact number of incongruent solutions is g . In the present case $g = n$, and the first congruence in (17) has a solution, by Theorem 1, so that $\sigma^{(p-1)/n} \equiv 1 \pmod{p}$. Also, $\sigma_1^{p-1} \equiv 1 \pmod{p}$ so that

$$\sigma^{(p-1)/n} \sigma_1^{p-1} = (\sigma\sigma_1^n)^{(p-1)/n} \equiv 1 \pmod{p},$$

and the lemma is proved.

We shall consider modules of the form

$$(19) \quad M = Z_0 + y\tau_1^{-1}Z_0 + \cdots + y^{n-1}\tau_{n-1}^{-1}Z_0$$

where

$$(20) \quad y = (\lambda - u)\beta^{-1}$$

with λ satisfying (17), β given by (15), and where the τ_i are rational integers such that

$$(21) \quad \tau_{n-1} \text{ divides } \sigma, \quad \tau_i \text{ divides } \tau_{i+1}$$

for $i = 1, \dots, n-2$. The τ_i will be chosen so that M is a ring. First, for any a_0 in Z_0 we find by a simple computation that $a_0 y = (a_0 - a_0^s)\lambda\beta^{-1} + ya_0^s$. The ramification order of p in Z over R is n so that the inertial group of p in Z over R is the complete galois group of Z over R . Hence $a_0 \equiv a_0^s \pmod{\beta}$ and we have $a_0 y = ya_0^s + a_1 \lambda$, (a_1 in Z_0). A simple induction then yields

LEMMA 5. For every a_0 in Z_0 and every integer $i > 0$ we have

$$a_0 y^i = y^i a_0^{s^i} + y^{i-1} a_1 \lambda + \cdots + a_i \lambda^i, \quad a_i \text{ in } Z_0.$$

By means of an n -rowed matrix representation of D it may be verified† that the characteristic function of y is

$$(22) \quad t^n - \lambda \delta_1 t^{n-1} + \lambda^2 \delta_2 t^{n-2} - \cdots + (-\lambda)^{n-1} \delta_{n-1} t + (-1)^n \delta_n$$

where δ_n is the rational integer $\delta_n = (\lambda^n - \sigma)p^{-1}$ and, for $i < n$, δ_i is the i th elementary symmetric function of β^{-1} and its conjugates. The i th elementary symmetric function of the algebraic integer $p\beta^{-1}$ and its conjugates in Z is $p^i \delta_i$ which must then be a rational integer. Since $p^i \delta_i$ is divisible by

$$P^{i(n-1)} = P^{(i-1)n+n-i} = (p^{i-1})P^{n-i},$$

it follows that $p\delta_i$ is a rational integer divisible by P^{n-i} and hence by p when $i < n$. This proves that all of the coefficients of (22) are rational integers.

* L. E. Dickson, *Introduction to the Theory of Numbers*, 1931, p. 31, exercise 5.

† See H, p. 525.

Observe that the coefficient δ_n in (22) has the property that $\delta_n \sigma^{-1}$ is an integer prime to σ . An induction based on (22) yields

LEMMA 6. For $k=0, 1, \dots, n-2$ we have

$$y^{n+k} = y^{n-1}a_1 + y^{n-2}a_2 + \dots + a_n$$

with rational integral coefficients a_j such that

$$\begin{aligned} a_j &\equiv 0 \pmod{\lambda^{k+j}}, & j &= 1, \dots, n-k-1, \\ a_{n-k} &\equiv 0 \pmod{\sigma}, & (a_{n-k}\sigma^{-1}, \sigma) &= 1, \end{aligned}$$

and, if $k > 0$,

$$a_j \equiv 0 \pmod{\lambda^{k+j+1-n}}, \quad j = n-k+1, \dots, n.$$

Thus every a_j is divisible by σ .

The module $M = Z_0 + \sum_{i=1}^{n-1} y^i \tau_i^{-1} Z_0$ of (19) contains the set MZ_0 , that is, all sums of products aa_0 with a in M and a_0 in Z_0 . By Lemma 5, (21), and the fact that λ is divisible by σ , we see also that the sets $Z_0 y^i \tau_i^{-1}$ are all contained in M so that $Z_0 y^i \tau_i^{-1} Z_0 \leq M$, $Z_0 M \leq M$. Thus M is a ring if and only if we have

$$(23) \quad y^i \tau_i^{-1} M \leq M, \quad i = 1, \dots, n-1.$$

This is equivalent to the condition

$$(24) \quad y^i \tau_i^{-1} y^j \tau_j^{-1} = y^{i+j} (\tau_i \tau_j)^{-1} \text{ in } M, \quad i, j = 1, \dots, n-1.$$

When $i+j < n$, the condition (24) holds if and only if $\tau_i \tau_j$ divides τ_{i+j} . Otherwise $i+j = n+k$, ($k=0, 1, \dots, n-2$), and, by Lemma 6, (24) holds if and only if $\tau_i \tau_j$ divides each quantity $a_r \tau_{n-r}$, ($r=1, \dots, n$), where we define $\tau_0=1$. In particular, it is sufficient to have

$$(25) \quad \sigma^{k+r} \tau_{n-r} \equiv 0 \pmod{\tau_i \tau_j}, \quad r = 1, \dots, n-k-1,$$

$$(26) \quad \sigma^{k+r+1-n} \tau_{n-r} \equiv 0 \pmod{\tau_i \tau_j}, \quad r = n-k, \dots, n.$$

Since $\tau_i \tau_j$ divides σ^2 , (25) holds when $k+r \geq 2$. Otherwise $r=1$, $k=0$, and (25) becomes $\sigma \tau_{n-1} \equiv 0 \pmod{\tau_i \tau_j}$ which by (21) is satisfied. In (26) we have $k+r \geq n$, and see that the condition is not restrictive when $k+r > n$, $k+r+1-n \geq 2$. We have proved

LEMMA 7. Sufficient conditions that the module M of (19) be a ring are given by the following congruences:

$$(27) \quad \tau_{i+j} \equiv 0 \pmod{\tau_i \tau_j}, \quad i+j < n,$$

$$(28) \quad \sigma \tau_{i+j-n} \equiv 0 \pmod{\tau_i \tau_j}, \quad i+j \geq n.$$

Let us now make the definition $\tau_0=1$,

$$(29) \quad \tau_j = \prod_{i=1}^s q_i^{e_i}, \quad e_i = \left[\frac{j}{n_i} \right]$$

for $j=1, \dots, n-1$, and verify that this choice of the τ_j satisfies the conditions* of Lemma 7. The quantity $\tau_a \tau_b$ is exactly divisible by q_i^{e+f} , $e = [a/n_i]$, $f = [b/n_i]$. If $a+b < n$, the quantity τ_{a+b} is exactly divisible by q_i^g , $g = [(a+b)/n_i] \geq e+f$, so that (27) holds. If $a+b \geq n$, then τ_{a+b-n} has the exact factor q_i^g ,

$$g = \left[\frac{a+b-n}{n_i} \right] = \left[\frac{a+b}{n_i} \right] - \frac{n}{n_i} \geq e + f - \frac{n}{n_i}.$$

But σ has the factor q_i^{n/n_i} , $\sigma \tau_{a+b-n}$ has $q_i^{g+n/n_i} \geq q_i^{e+f}$ as factor, so that (28) holds. We have proved that M is a ring.

The ring M is a linear set of finite order over the domain of all rational integers; it contains Z_0 and hence all rational integers; and it contains $u = \lambda - \gamma\beta$ and hence a basis $u^{i-1}z_j$, ($i, j=1, \dots, n$), of D where the z_j form any integral basis of Z . These properties imply† that the quantities of M are all integral and that M is an order of D . This order is maximal in D if and only if‡ its discriminant is the discriminant§

$$(30) \quad \prod_{i=1}^s q_i^{n^2(n_i-1)/n_i}$$

of the algebra D .

The sets M and I have respective bases w and v given by the vectors

$$w = (z_1, \dots, z_n, \tau_1^{-1} y z_1, \dots, \tau_1^{-1} y z_n, \dots, \tau_{n-1}^{-1} y^{n-1} z_n) = (w_1, \dots, w_{n^2}),$$

$$v = (z_1, \dots, z_n, u z_1, \dots, u z_n, \dots, u^{n-1} z_n) = (v_1, \dots, v_{n^2}),$$

where the z_j form an integral basis of Z . There is a nonsingular matrix B with rational elements such that $w = vB$, and the discriminant of M is then

$$\Delta(w) = |T(w_i w_j)| = \Delta(v) \cdot |B|^2.$$

Here $\Delta(v)$ is the discriminant $|T(v_i v_j)|$ of the basis v , and $\Delta(v) = (\sigma p)^{n(n-1)}$. To compute $|B|^2$ we observe¶ that when the matrix B is expressed as an

* Note that this choice of the τ_j makes $\tau_1, \dots, \tau_{n-1}$ prime to q_i . Hence $\tau_1 = \dots = \tau_{n-1} = 1$.

† Deuring, op. cit., p. 71, Theorem 9.

‡ E. Artin, *Zur Arithmetik hyperkomplexer Zahlen*, Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität, vol. 5 (1928), p. 265.

§ Reichardt, *Die Diskriminante einer normalen einfachen Algebra*, Journal für die reine und angewandte Mathematik, vol. 173 (1935), pp. 31-34.

¶ See H, p. 523.

¶ Ibid., p. 526.

n -rowed matrix whose elements are $n \times n$ matrices B_{ij} , then every B_{ij} below the main diagonal is a zero matrix, B_{11} is an identity matrix, and every matrix B_{jj} , ($j > 1$), has determinant equal, except possibly for sign, to the norm

$$[N(\tau_{j-1}\beta\beta^S \cdots \beta^{S^{j-2}})]^{-1} = (\tau_{j-1}p^{j-1})^{-1}.$$

Then $|B|^2 = |B_{11}B_{22} \cdots B_{nn}|^2$ has the value

$$|B|^2 = (\tau_1 \cdots \tau_{n-1})^{-2n} p^{-n(n-1)}$$

so that $\Delta(w) = \sigma^{n(n-1)}(\tau_1 \cdots \tau_{n-1})^{-2n}$. But

$$\tau_1 \cdots \tau_{n-1} = \prod_{i=1}^n q_i^{[(n/n_i-1)+(n/n_i-2)+\cdots+1]n_i} = \prod q_i^{(n/n_i-1)n/2}$$

and $\Delta(w) = \prod q_i^{n^2(n_i-1)/n_i}$ which is the formula (30) for the discriminant of D . Thus M is a maximal order of D .

THEOREM 3. *Let D be an algebra of Theorem 1 with normalized cyclic generation (Z, S, σ) as described in that theorem, $D = Z + uZ + \cdots + u^{n-1}Z$, $u^n = \sigma$. Then n distinct maximal orders in D are given by the modules*

$$M(\lambda) = Z_0 + y\tau_1^{-1}Z_0 + \cdots + y^{n-1}\tau_{n-1}^{-1}Z_0,$$

where Z_0 is the maximal order of Z , the τ_i are rational integers defined by (29), and $y = (\lambda - u)\beta^{-1}$, with β defined by (15) and λ varying over the n rational integers defined by (17). Each $M(\lambda)$ contains the order

$$I = Z_0 + uZ_0 + \cdots + u^{n-1}Z_0$$

associated with the cyclic generation (Z, S, σ) of D .

That $M(\lambda_1)$ is distinct from $M(\lambda_2)$ was proved in H, p. 527, by showing that the corresponding quantities $y = y_1$, $y = y_2$ are such that $y_1 - y_2$ is not integral.

5. Maximal orders in direct products. In view of the factorization of any normal division algebra D into a direct product of normal division algebras D_i whose degrees are powers of distinct primes, we may inquire whether maximal orders of D can be obtained simply in terms of those of the D_i . We shall solve this problem under certain hypotheses on the D_i and shall obtain some further results on the general problem.

Let A_1 and A_2 be cyclic algebras of relatively prime degrees over R and $A = A_1 \times A_2$. A ramification spot of A must be a ramification spot for one of the A_i . Conversely, suppose that one of the A_i does not split at q . Then A_{1q} and A_{2q} have indices d_1 and d_2 which are relatively prime and one of which

is greater than unity. Hence A_q has index $d_1 d_2 > 1$. Thus the ramification spots of A are those of A_1 together with those of A_2 .

If A_1 and A_2 have cyclic generation fields Z_1 and Z_2 , respectively, then A has the cyclic generation field $Z_1 \times Z_2$. This fact will be used several times in this section and may be verified by a direct computation and also, for algebras over R , in the following way.

LEMMA 8. Let $A_i = (Z_i, S_i, \sigma_i)$ be a cyclic algebra of degree m_i over R , ($i=1, 2$), where $(m_1, m_2)=1$. Then $A = A_1 \times A_2$ has a cyclic generation $A = (Z, S, \sigma)$ where $Z = Z_1 \times Z_2$, $S = S_1 S_2$, $\sigma = \sigma_1^{m_2} \sigma_2^{m_1}$.

The composite of the Z_i is their direct product, so that (Z, S, σ) has degree $m_1 m_2$ over R . If the invariants of A_i are denoted by ν_{iq} for every prime spot q and those of A by ν_q , then*

$$\nu_q \equiv m_2 \nu_{1q} + m_1 \nu_{2q} \pmod{m_1 m_2}.$$

We have

$$\begin{aligned} (\sigma, Z | q) &= \prod_{i,j} (\sigma_i, Z_j | q)^{m_1 m_2 / m_i} = \prod_i (\sigma_i, Z_i | q)^{m_1 m_2 / m_i} \\ &= S_1^{m_2 \nu_{1q}} S_2^{m_1 \nu_{2q}} = (S_1 S_2)^{m_2 \nu_{1q} + m_1 \nu_{2q}} = S^{\nu_q}. \end{aligned}$$

Hence (Z, S, σ) has the same invariants ν_q and degree $m_1 m_2$ as A . Thus the lemma is proved.

Let J_1 and J_2 be any orders in A_1 and A_2 , respectively, and consider the product $J_1 J_2$ in $A_1 \times A_2$, consisting of all sums of products $a_1 a_2$ with a_i in J_i . The set $J = J_1 J_2$ is an order in A as one can easily verify. If bases of J_1 and J_2 over the rational integers are given by (u_1, \dots, u_{m_1}) and (v_1, \dots, v_{m_2}) , respectively, $J_1 J_2$ has a basis $(u_1 v_1, \dots, u_i v_j, \dots, u_{m_1} v_{m_2})$.

LEMMA 9. If J_i has discriminant Δ_i , ($i=1, 2$), then $J = J_1 J_2$ has discriminant $\Delta_0 = \Delta_1^{m_2} \Delta_2^{m_1}$.

The basis given above for J may be designated by $(w_1, \dots, w_{m_1 m_2})$, and then $\Delta_0 = |T(w_i w_j)|$ where T is the trace function in A . Let T_i be the trace in A_i , and let a_i be any element of A_i . We shall show that $T(a_1 a_2) = T_1(a_1) T_2(a_2)$.

Let W_i be a basis of A_i relative to a cyclic generation field Z_i of A_i for $i=1, 2$. Then the equation of $a_i W_i = W_i B_i$ defines a set of matrices B_i , with elements in Z_i , forming an algebra equivalent to A_i under the correspondence $a_i \longleftrightarrow B_i$ for every a_i of A_i , and $T_i(a_i)$ is defined to be the trace of the matrix B_i . Since m_1 and m_2 are relatively prime, the composite $Z = Z_1 \times Z_2$ is a cyclic generation field of A , and a basis of A relative to Z is given by the vector W

* Hasse, *Theory of cyclic algebras over an algebraic number field*, loc. cit., p. 179, Theorem 4.

consisting of the products of each of the elements of W_1 by each of W_2 . Then $aW = WB$ defines a representation $a \mapsto B$ of A , and $T(a)$ is the trace of the matrix B . We write $W_i = (w_{i1}, \dots, w_{im_i})$ and have

$$\begin{aligned} a_1 w_{1r} &= \sum_f w_{1f} b_{1fr}, & a_2 w_{2t} &= \sum_g w_{2g} b_{2gt}, \\ a_1 w_{1r} a_2 w_{2t} &= a_1 a_2 w_{1r} w_{2t} = \sum_{f,g} w_{1f} w_{2g} b_{1fr} b_{2gt}. \end{aligned}$$

Hence the matrix B corresponding to $a = a_1 a_2$ has elements $b_{1fr} b_{2gt}$ and has, as desired, the trace

$$T(a_1 a_2) = \sum_{r,t} b_{1rr} b_{2tt} = \left(\sum_r b_{1rr} \right) \left(\sum_t b_{2tt} \right) = T_1(a_1) T_2(a_2).$$

Since $w_x w_y = u_i v_h u_j v_k$, we may write $T(w_x w_y) = T(u_i u_j v_h v_k) = T_1(u_i u_j) T_2(v_h v_k)$. Consider the matrices $C_1 = (T_1(u_i u_j))$ and $C_2 = (T_2(v_h v_k)) = (c_{hk})$. The discriminant $|T(w_x w_y)|$ of J is the determinant $\Delta_0 = |C_1 c_{hk}|$ of a matrix which we have written as a square matrix of $m_2^2 = k_2$ rows whose elements are square matrices of $m_1^2 = k_1$ rows. When C_2 is one-rowed, we have $|C_1 c_{hk}| = |C_1|^{k_2} |C_2|^{k_1}$ since then $k_2 = 1$, and we now assume that this formula holds for all matrices C_2 of $k_2 - 1$ rows. We may assume $c_{11} \neq 0$ and then may replace the blocks $C_1 c_{h1}$ by zero matrices under elementary transformations which replace the blocks $C_1 c_{hk}$ by $C_1 d_{hk}$, $d_{hk} = c_{hk} - c_{h1} c_{1k} c_{11}^{-1}$. In the remainder of this paragraph the subscripts h and k on c_{hk} will vary over $1, \dots, k_2$ and those on d_{hk} will vary over $2, \dots, k_2$. We have

$$\Delta_0 = |C_1 c_{hk}| = |C_1 c_{11}| \cdot |C_1 d_{hk}| = |C_1| \cdot |c_{11}|^{k_1} \cdot |C_1|^{k_2-1} \cdot |d_{hk}|^{k_1}$$

by our induction. But $|c_{hk}| = |c_{11}| |d_{hk}|$ so that $\Delta_0 = |C_1|^{k_2} |C_2|^{k_1}$, and the lemma is proved.

The discriminant of A is the product*

$$\Delta = \prod_q q^{e_q}, \quad e_q = (n_q - 1)n^2/n_q,$$

where q varies over all ramification spots of A , n is the degree $m_1 m_2$ of A , and n_q is the q -index of A . Then $n_q = m_{1q} m_{2q}$ where m_{iq} is the q -index of A_i . Let Δ_i be the discriminant of A_i . Then a direct computation shows that if A_1 and A_2 have no ramification spots in common, the discriminant of A is $\Delta_1^{m_2^2} \Delta_2^{m_1^2}$. Otherwise the Δ_i have common factors q , and in fact we find that in general A has discriminant

$$\Delta = \Delta_1^{m_1^2} \Delta_2^{m_2^2} \prod_q q^{-(m_{1q}-1)(m_{2q}-1)n^2/n_q}$$

* Reichardt, op. cit.

where the product is taken over all common ramification spots q of A_1 and A_2 . An immediate consequence of this formula and Lemma 9 is stated now.

THEOREM 4. *Let A_1 and A_2 be normal simple algebras of relatively prime degrees m_1 and m_2 over R , and let M_1 and M_2 be any maximal orders in A_1 and A_2 , respectively. Then M_1M_2 is a maximal order in $A = A_1 \times A_2$ if and only if the discriminants Δ_1 and Δ_2 of A_1 and A_2 are relatively prime. In this case the discriminant of A is $\Delta_1^{m_2} \Delta_2^{m_1}$.*

This is an analogue of a known theorem* on algebraic fields over R with relatively prime discriminants. That $M = M_1M_2$ is maximal may also be proved by using Hasse's determination† of all maximal orders in the q -adic algebra A_q . We show by this means that for every prime spot q the q -component M_q is a maximal order of A_q . But this is a necessary and sufficient condition that M be maximal in A .

An application of Lemma 8 and Theorem 1 yields the following result which may be useful in the determination of maximal orders in a direct product.

THEOREM 5. *Let D be a direct product $D_1 \times \cdots \times D_t$ of normal division algebras D_i of Theorem 1 such that the degrees m_i of the D_i are relatively prime in pairs, and let $n = m_1 \cdots m_t$. Then each D_i has a cyclic generation (Z_i, S_i, σ_i) as described in Theorem 1, and D has a cyclic generation*

$$D = (Z, S, \sigma), \quad Z = Z_1 \times \cdots \times Z_t, \quad S = S_1 \cdots S_t, \quad \sigma = \prod_i \sigma_i^{n/m_i}.$$

The generations of the D_i may be chosen so that the conductors p_1, \cdots, p_t of Z_1, \cdots, Z_t are distinct primes, and are not ramification spots of D . The former property implies that the maximal order Z_0 of Z is the product of the maximal orders Z_{0i} of the fields Z_i .

* D. Hilbert, *Gesammelte Abhandlungen*, vol. 1, 1932, p. 146. The result of Theorem 4 was also obtained in a different way by K. Shoda and T. Nakamura in the paper *Über das Produkt zweier Algebrenklassen mit zueinander primen Diskriminanten*, Proceedings of the Imperial Academy of Japan, vol. 10 (1934), pp. 443-446.

† H. Hasse, *Über p -adische Schiefkörper und ihre Bedeutung für die Arithmetik hyperkomplexer Zahlssysteme*, Mathematische Annalen, vol. 104 (1931), pp. 495-534, Theorem 47.

CONVERGENCE PROPERTIES OF ANALYTIC FUNCTIONS OF FOURIER-STIELTJES TRANSFORMS*

BY

ROBERT H. CAMERON AND NORBERT WIENER

1. **Introduction.** Wiener and Pitt† have given conditions under which the reciprocal of an absolutely convergent Fourier-Stieltjes integral is again an absolutely convergent Fourier-Stieltjes integral. It is the purpose of this paper to generalize this result in two directions. We replace reciprocals by general analytic functions which may even be multiple-valued, and we replace absolute convergence by finiteness of certain more general norms. These norms are of two types, both of which depend on a parameter θ , ($0 < \theta \leq 1$); and both reduce to total variation when $\theta = 1$.

Following the notation of (WP), we let $f(x)$ denote a function of bounded variation in $(-\infty, \infty)$ for which $2f(x) = f(x+0) + f(x-0)$, and let

$$F(x) = \int_{-\infty}^{\infty} e^{-iyx} df(y).$$

We write

$$f(x) = h(x) + g(x) + s(x),$$

where $h(x)$ is a step-function, $g(x)$ is absolutely continuous, and $s(x)$ is continuous and has a zero derivative almost everywhere. We refer to $h(x)$, $g(x)$, $s(x)$ as the discrete, smooth, and singular parts of $f(x)$, and to their Fourier-Stieltjes transforms $H(x)$, $G(x)$, $S(x)$ as the almost periodic, transient, and unpredictable parts of $F(x)$; of course we have $F(x) = H(x) + G(x) + S(x)$. Moreover $h(x)$, $g(x)$, $s(x)$ are each of bounded variation and essentially uniquely determined by $f(x)$, while $H(x)$, $G(x)$, $S(x)$ are uniquely determined by $F(x)$.

We define for $0 < \theta \leq 1$,

$$T_{\theta}^* \{F(x)\} = 2 \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} |df(y)| \right]^{\theta}$$

and‡

* Presented to the Society, December 28, 1938; received by the editors January 21, 1939.

† On absolutely convergent Fourier-Stieltjes transforms, Duke Mathematical Journal, vol. 4 (1938), pp. 420-436. This paper will be referred to as (WP).

‡ We use the symbol $\int |dh(y)|^{\theta}$ to mean the sum of the θ powers of the jumps of $h(y)$. By a jump we mean the whole jump $|h(y+0) - h(y-0)|$, not a half jump $|h(y+0) - h(y)|$.

$$T_{\theta}^{**}\{F(x)\} = 2 \int_{-\infty}^{\infty} |dh(y)|^{\theta} + 2 \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} |dg(y) + ds(y)| \right]^{\theta};$$

and we say that $F(x) \in A_{\theta}^*$ or $F(x) \in A_{\theta}^{**}$ if $T_{\theta}^*[F(x)] < \infty$ or $T_{\theta}^{**}[F(x)] < \infty$. We use the symbol T_{θ} to stand for T_{θ}^* or T_{θ}^{**} as \pm stands for $+$ or $-$; and similarly we use A_{θ} for A_{θ}^* or A_{θ}^{**} . Thus, the previous statement might have been written $F(x) \in A_{\theta}$ if $T_{\theta}[F(x)] < \infty$. We will also suppress the θ when no confusion will be caused.

We now state the main theorem of this paper:

THEOREM I. Let $F(x) \in A_{\theta}$, let R be the closure of the set of values of $F(x)$, and let R^* be the set of complex numbers whose distance from R is not greater than $\{T_{\theta}[S(x)]\}^{1/\theta}$. Let $\mathcal{F}(z)$ be a multiple-valued function defined on an open set \mathcal{R} containing R^* ; and let $\mathcal{F}(z)$ consist of exactly n distinct nonintersecting analytic sheets in the neighborhood of each point of \mathcal{R} . Let the n continuous branches of $\mathcal{F}[F(x)]$ and $\mathcal{F}[H(x)]$ be denoted in some arbitrary order by $[\mathcal{F}(F(x))]_j$, ($j=1, \dots, n$), and by $[\mathcal{F}(H(x))]_j$, ($j=1, \dots, n$). Then there exist two permutations p_1, \dots, p_n and p'_1, \dots, p'_n (each unique) of the numbers $1, 2, \dots, n$ such that

$$(1.1) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \int_0^N |[\mathcal{F}(F(x))]_j - [\mathcal{F}(H(x))]_{p_j}|^2 dx = 0,$$

$$(1.2) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \int_{-N}^0 |[\mathcal{F}(F(x))]_j - [\mathcal{F}(H(x))]_{p'_j}|^2 dx = 0.$$

Moreover if for any particular j we have $p_j = p'_j$, then $[\mathcal{F}(F(x))]_j \in A_{\theta}$.

2. Properties of the norms. We must first establish the fact that the norms $T_{\theta}\{F(x)\}$ satisfy the axioms

- I. $T(F_1) + T(F_2) \geq T(F_1 + F_2)$;
- II. $T(F_1)T(F_2) \geq T(F_1 F_2)$;
- III. $|a|^{\theta} T_{\theta}(F) = T_{\theta}(aF)$.

The first of these relations follows immediately from the inequality $a^{\theta} + b^{\theta} \geq (a+b)^{\theta}$, which holds whenever $a \geq 0$, $b \geq 0$, and $0 < \theta \leq 1$ (as we can readily see by choosing $a > b$ and considering the function $(b/a)^{\theta} + 1 - (b/a + 1)^{\theta}$ and its θ derivative). The third relation is obvious; and it therefore only remains to prove axiom II. Assuming therefore that $F(x) = F_1(x)F_2(x)$, we obtain

$$f(y) = \int_{-\infty}^{\infty} f_1(y-u)df_2(u)$$

except at a countable set of points. Then

$$\begin{aligned}
 T_{\theta}^* \{F_1(x)F_2(x)\} &= 2 \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} \left| d_y \int_{-\infty}^{\infty} f_1(y-u)df_2(u) \right| \right]^{\theta} \\
 &= 2 \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} \left| d_y \sum_{m=-\infty}^{\infty} \int_m^{m+1} f_1(y-u)df_2(u) \right| \right]^{\theta} \\
 &\leq 2 \sum_{n=-\infty}^{\infty} \left[\sum_{m=-\infty}^{\infty} \int_n^{n+1} \left| d_y \int_m^{m+1} f_1(y-u)df_2(u) \right| \right]^{\theta} \\
 &\leq 2 \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \left[\int_m^{m+1} |df_2(u)| \int_{n-m-1}^{n-m+1} |df_1(u)| \right]^{\theta} \\
 &= 2 \sum_{m=-\infty}^{\infty} \left[\int_m^{m+1} |df_2(u)| \right]^{\theta} \sum_{n=-\infty}^{\infty} \left[\int_{n-1}^{n+1} |df_1(y)| \right]^{\theta} \\
 &\leq T_{\theta}^* [F_2(x)] \cdot T_{\theta}^* [F_1(x)].
 \end{aligned}$$

Thus II holds for T_{θ}^* ; and since T_{θ}^* and T_{θ}^{**} are identical for functions with zero almost periodic part, II holds for T_{θ}^{**} applied to functions of the form $F(x) = G(x) + S(x)$. Since I holds for T_{θ}^{**} , we need merely show that II holds if $F_1(x) = H_1(x)$ and $F_2(x) = H_2(x)$ and also holds if $F_1(x) = H_1(x)$ and $F_2(x) = G_2(x) + S_2(x)$. But $H(x)$ is merely an infinite sum of terms of the form $ae^{i\lambda x}$, and since I can be extended to infinite sums, we need merely show that II holds for products of the form $a_1e^{i\lambda_1 x}a_2e^{i\lambda_2 x}$ and $ae^{i\lambda x}[G(x) + S(x)]$. Direct substitution takes care of the first of these products; and the proof is completed by noting that if m is the greatest integer less than λ ,

$$\begin{aligned}
 T_{\theta}^{**} \{ae^{i\lambda x}[G(x) + S(x)]\} &= T_{\theta}^{**} \left\{ a \int_{-\infty}^{+\infty} e^{ixy} d[g(y-\lambda) + ds(y-\lambda)] \right\} \\
 &= 2 |a|^{\theta} \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} |dg(y-\lambda) + ds(y-\lambda)| \right]^{\theta} \\
 &\leq 2 |a|^{\theta} \sum_{n=-\infty}^{\infty} \left[\int_{n-m-1}^{n-m+1} |dg(y) + ds(y)| \right]^{\theta} \\
 &\leq T_{\theta}^{**} \{ae^{i\lambda x}\} \cdot T_{\theta}^{**} \{G(x) + S(x)\}.
 \end{aligned}$$

3. Functions of small norm. We begin to prove Theorem I by first proving that the special case of it in which $\mathcal{F}(z)$ is single-valued and $F(x)$ is a constant plus a function of small norm.

LEMMA 1. Let $F(x) \in A_{\theta}$, let $F(x) = \tau + F_1(x)$ where $T_{\theta}[F_1(x)] < K^{\theta}$, and let $\mathcal{F}(z)$ be analytic in a circle about τ of radius K . Then $\mathcal{F}\{F(x)\} \in A_{\theta}$.

For $\mathcal{F}(z)$ has a Taylor's series $\mathcal{F}(z) = \sum_{n=0}^{\infty} a_n(z-\tau)^n$ converging when

$|z - \tau| < K$. Then if $T_\theta[F_1(x)] < K_1^\theta < K^\theta$, $\sum_{n=0}^\infty |a_n| K_1^n$ converges and $|a_n| K_1^n$ is bounded in n . Hence $\sum_{n=0}^\infty |a_n|^\theta W^n$ converges when $0 < W < K_1^\theta$. Thus

$$\begin{aligned} T_\theta[\mathcal{F}(F(x))] &= T_\theta \left[\sum_{n=0}^\infty a_n(F(x) - \tau)^n \right] \\ &= T_\theta \left[\sum_{n=0}^\infty a_n(F_1(x))^n \right] \leq \sum_{n=0}^\infty T_\theta [a_n(F_1(x))^n] \\ &\leq \sum_{n=0}^\infty |a_n|^\theta \{T_\theta[F_1(x)]\}^n, \end{aligned}$$

and since $T_\theta[F_1(x)] < K_1^\theta$, it follows that the last sum is finite. Thus $\mathcal{F}(F(x)) \in A_\theta$.

4. The space $C\mathfrak{T}_n$. Let \mathfrak{T}_n be the set of points each of which consists of n ordered numbers, each reduced modulo 2π . Let C be the set of all real numbers, together with one special symbol ∞ . Let $C\mathfrak{T}_n$ be the product space of C and \mathfrak{T}_n .

The set of points (x_1, \dots, x_n) of \mathfrak{T}_n which satisfy

$$|x_j - x'_j| < \epsilon \pmod{2\pi}, \quad j = 1, \dots, n,$$

is called the ϵ -neighborhood of (x'_1, \dots, x'_n) . The set of points x of C which satisfy $|x - x'| < \epsilon$ is called the ϵ -neighborhood of x . The set of finite points x which satisfy $|x| > 1/\epsilon$ together with ∞ , is called the ϵ -neighborhood of ∞ . Product neighborhoods such as the ϵ -neighborhood of $(x_1, \dots, x_n; x)$ or $(x_1, \dots, x_n; \infty)$ are defined in the usual way. The ϵ -neighborhood of $(x_1, \dots, x_n; \infty)$ will be called an infinite $C\mathfrak{T}_n$ neighborhood, and $(x_1, \dots, x_n; \infty)$ will be called an infinite point of $C\mathfrak{T}_n$. It is obvious that the Heine-Borel theorem holds for the whole space.

5. Finiteness of the norm a local property. A function $F(x)$ is called locally of finite norm in a finite C neighborhood N if there exists a function $F^*(x)$ which is of finite norm and equals $F(x)$ when x is in N . A function $F(x)$ is called locally of finite norm with respect to $\lambda_1, \dots, \lambda_n$ in a $C\mathfrak{T}_n$ neighborhood N if there exists a function $F^*(x)$ which is of finite norm and equals $F(x)$ when $(\lambda_1 x, \dots, \lambda_n x; x)$ is in N .

LEMMA 2. Let $\lambda_1, \dots, \lambda_n$ be given. Then a necessary and sufficient condition that a function $f(x)$ be of finite norm is that it be locally of finite norm in a C neighborhood of each finite point of C and locally of finite norm with respect to $\lambda_1, \dots, \lambda_n$ in a $C\mathfrak{T}_n$ neighborhood of each infinite point of $C\mathfrak{T}_n$.

The necessity of the condition is obvious; so we need only prove sufficiency. We note at the outset that the hypothesis implies that $f(x)$ is locally of finite norm with respect to $\lambda_1, \dots, \lambda_n$ in a $C\mathfrak{T}_n$ neighborhood of every point

of $C\mathfrak{C}_n$. For a function which equals $f(x)$ in the ϵ -neighborhood of the point x_0 of C necessarily equals $f(x)$ when $(\lambda_1 x, \dots, \lambda_n x; x)$ is in the ϵ -neighborhood of the point $(x_1, \dots, x_n; x_0)$ of $C\mathfrak{C}_n$. Thus for each point P of $C\mathfrak{C}_n$ there is an $\epsilon_P > 0$ and a function $f_P(x)$ which is of finite norm and which equals $f(x)$ when $(\lambda_1 x, \dots, \lambda_n x; x)$ is in the ϵ_P -neighborhood N_P of P . Then by the Heine-Borel theorem there is a finite number of points P_1, \dots, P_q , such that the $\epsilon/2$ -neighborhoods of P_i cover $C\mathfrak{C}_n$. Choose an integer

$$N > 2\pi [\min (\epsilon_{P_1}, \dots, \epsilon_{P_q})]^{-1} \max [|\lambda_1|, \dots, |\lambda_n|; 1].$$

Let $\Phi^*(x)$ be an even function which is zero on $|x| > 1$, unity at $x=0$, and is continuous and has derivatives of all orders everywhere and satisfies $\Phi^*(x) + \Phi^*(1-x) = 1$ on $0 \leq x \leq 1$. Thus, to be specific, we may define

$$\Phi^*(x) = \begin{cases} \frac{1}{2} - \frac{\int_0^{2|x|-1} e^{(\xi^2-1)^{-1}} d\xi}{2 \int_0^1 e^{(\xi^2-1)^{-1}} d\xi} & \text{if } 0 < |x| \leq 1, \\ 0 & \text{if } 1 \leq |x|. \end{cases}$$

We obviously have

$$\sum_{k=-N^2}^{N^2} \Phi(x-k) = \begin{cases} 1 & \text{if } |x| \leq N^2, \\ 0 & \text{if } |x| > (N+1)^2. \end{cases}$$

Moreover if

$$\begin{cases} \Phi(x) = \Phi^*\left(\frac{Nx}{\pi}\right), & |x| < \pi, \\ \Phi(x) = \Phi(x+2\pi), & \text{for all } x, \end{cases}$$

then

$$\sum_{k=1}^{2N} \Phi\left(x - \frac{\pi k}{N}\right) = 1, \quad \text{for all } x.$$

Thus

$$\sum_{k_1, \dots, k_n=1}^{2N} \Phi\left(\lambda_1 x - \frac{\pi}{N} k_1\right) \cdots \Phi\left(\lambda_n x - \frac{\pi}{N} k_n\right) = 1 \quad \text{for all } x,$$

and

$$\begin{aligned} f(x) &= \sum_{k=-N^2}^{N^2} \Phi^*(Nx-k) f(x) \\ &+ \left[1 - \sum_{k=-N^2}^{N^2} \Phi^*(Nx-k) \right] \sum_{k_1, \dots, k_n=1}^{2N} \Phi\left(\lambda_1 x - \frac{\pi}{N} k_1\right) \\ &\cdots \Phi\left(\lambda_n x - \frac{\pi}{N} k_n\right) f(x) \end{aligned}$$

for all x . Thus if we show that $\Phi^*(Nx-k)f(x)$ and

$$\left[1 - \sum_{k=-N^2}^{N^2} \Phi^*(Nx-k)\right] \Phi\left(\lambda_1 x - \frac{\pi}{N} k_1\right) \cdots \Phi\left(\lambda_n x - \frac{\pi}{N} k_n\right) f(x)$$

are of finite norm for all k, k_1, \dots, k_n , it follows that $f(x)$ is of finite norm.

To show this for $\Phi^*(Nx-k)$, consider the point P_i whose $\epsilon_{P_i}/2$ -neighborhood covers

$$\left(\frac{k}{N} \lambda_1, \dots, \frac{k}{N} \lambda_n; \frac{k}{N}\right).$$

The ϵ_{P_i} -neighborhood \mathcal{N}_i of this point covers the $\epsilon_{P_i}/2$ -neighborhood of

$$\left(\frac{k}{N} \lambda_1, \dots, \frac{k}{N} \lambda_n; \frac{k}{N}\right).$$

Thus when $|Nx-k| < 1$, $(\lambda_1 x, \lambda_2 x, \dots, \lambda_n x; x)$ is in \mathcal{N}_i , and when $\Phi^*(Nx-k)$ is not zero, $f_{P_i}(x)$ equals $f(x)$ and for all x

$$T^*(Nx-k)f(x) = T^*(Nx-k)f_{P_i}(x),$$

which is of finite norm.

Again, consider the point P_l whose $\epsilon_{P_l}/2$ -neighborhood covers the point $(k_1\pi/N, \dots, k_n\pi/N; \infty)$. The ϵ_{P_l} -neighborhood \mathcal{N}_l of P_l covers the $\epsilon_{P_l}/2$ -neighborhood of $(k_1\pi/N, \dots, k_n\pi/N; \infty)$. Thus when

$$|\lambda_1 x - k_1\pi/N| < \pi/N, \dots, |\lambda_n x - k_n\pi/N| < \pi/N, \quad |x| > N,$$

$(\lambda_1 x, \lambda_2 x, \dots, \lambda_n x, x)$ is in \mathcal{N}_l and

$$\begin{aligned} & \left[1 - \sum_{k=-N^2}^{N^2} \Phi^*(Nx-k)\right] \Phi\left(\lambda_1 x - \frac{\pi}{N} k_1\right) \cdots \Phi\left(\lambda_n x - \frac{\pi}{N} k_n\right) f(x) \\ &= \left[1 - \sum_{k=-N^2}^{N^2} \Phi^*(Nx-k)\right] \Phi\left(\lambda_1 x - \frac{\pi}{N} k_1\right) \cdots \Phi\left(\lambda_n x - \frac{\pi}{N} k_n\right) f_{P_l}(x). \end{aligned}$$

Thus $f(x)$ is of finite norm.

6. Periodic functions with assigned derivatives and small norm. Section 5 makes it necessary only to show that $\mathcal{Y}[F(x)]$ is locally of finite norm corresponding to all infinite points $C\mathfrak{T}_n$ and all finite points of C ; and §3 shows that this will be established if we show how to replace functions by others locally equivalent but of sufficiently small norm. We begin by replacing exponential functions by locally equivalent functions of small norm.

The first step is to find functions of small norm having derivatives at the origin equal to those of $e^{iz}-1$.

Such a function is given by

$$\Psi_{m,p}(x) = e^{(i/p)\psi_m(px)} - 1$$

where

$$(6.1) \quad \psi_m(x) = x - \frac{2\pi \int_0^x \sin^2 m \xi d\xi}{\int_0^{2\pi} \sin^2 m \xi d\xi};$$

for $\psi_m(x)$ is obviously periodic and hence is an exponential polynomial. Thus $\psi_m(px)$ has norms independent of p (for p an integer greater than 1) and the norms of $(1/p)\psi_m(px) = O(1/p^0)$. Thus for fixed m the norm of $\Psi_{m,p}(x)$ can be made arbitrarily small by making p sufficiently great; and since $\psi_m(x) = x + O(x^{2m+1})$ at the origin, it follows that $\Psi_{m,p}(x) = e^{ix} - 1 + O(x^{2m+1})$, and $\Psi_{m,p}(x)$ and $e^{ix} - 1$ have the same first $2m$ derivatives at $x=0$ and points congruent (mod 2π).

7. Locally exponential functions of small norm. Now to obtain a function of small norm which is actually equal to $e^{ix} - 1$ in the neighborhood of points congruent to zero (mod 2π) we introduce the function

$$\Omega(\epsilon, x) = \begin{cases} 1, & 0 \leq |x| \leq \epsilon \pmod{2\pi}, \\ \frac{1}{2} - \frac{\int_0^{|x|/\epsilon-3} e^{(\xi^2-1)^{-1}} d\xi}{2 \int_0^1 e^{(\xi^2-1)^{-1}} d\xi}, & \epsilon \leq |x| \leq 2\epsilon \pmod{2\pi}, \\ 0, & 2\epsilon \leq |x| \leq \pi \pmod{2\pi} \end{cases}$$

which is obviously of finite norm for all θ and of period 2π .

Let $H(x)$ be a periodic function of period 2π whose first m derivatives are continuous everywhere. Then if $H(0) = H'(0) = \dots = H^{(m)}(0) = 0$, and $1/m < \theta \leq 1$, we shall show that

$$(7.1) \quad \lim_{\epsilon \rightarrow 0} T_\theta \{H(x)\Omega(\epsilon, x)\} = 0.$$

For if $H(x)\Omega(\epsilon, x) = P(\epsilon, x) = \sum_{n=-\infty}^{\infty} p_n(\epsilon) e^{inx}$, we have by integration by parts

$$p_n = \frac{1}{2\pi} \int_0^{2\pi} P(\epsilon, x) e^{-inx} dx = \frac{1}{2\pi(in)^m} \int_0^{2\pi} P^{(m)}(\epsilon, x) e^{-inx} dx \quad \text{for } n \neq 0.$$

Thus if $L(\epsilon)$ is the greatest value taken on by either $|P(\epsilon, x)|$ or $|P^{(m)}(\epsilon, x)|$ for all x , we have $|p_n(\epsilon)| \leq 2L(\epsilon)/(1+|n|^m)$ for all n and

$$T_\theta \{P(\epsilon, x)\} \leq 2[L(\epsilon)]^\theta \sum_{n=-\infty}^{\infty} \frac{2}{(1+|n|^m)^\theta} \quad \text{for all } m.$$

Since $m\theta > 1$, this sum converges, and we need merely show that $L(\epsilon) \rightarrow 0$ to

establish (7.1). But $\max_x |P(\epsilon, x)| \rightarrow 0$ as $\epsilon \rightarrow 0$ since $H(x)$ is continuous and vanishes at zero and $\Omega(\epsilon, x)$ is bounded and is zero when $|x| \geq 2\epsilon \pmod{2\pi}$. Moreover

$$P^{(m)}(\epsilon, x) = \sum_{j=0}^m C_{m,j} H^{(m-j)}(x) \Omega^{(j)}(\epsilon, x),$$

and we shall show that each term of this sum approaches zero uniformly as $\epsilon \rightarrow 0$. In the interval $|x| < 2\epsilon$ the function $\Omega(\epsilon, x)$ is a function of x/ϵ alone, and hence its j 'th derivative with respect to x is less than $C_j \epsilon^{-j}$, where C_j is independent of ϵ and x . Moreover $H^{(m-j)}(x) x^{-j} \rightarrow 0$ as $x \rightarrow 0$; so

$$\max_{|x| \leq 2\epsilon} \frac{|H^{(m-j)}(x)|}{\epsilon^j} \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0.$$

Thus

$$\begin{aligned} \max_x |H^{(m-j)}(x) \Omega^{(j)}(\epsilon, x)| &= \max_{|x| \leq 2\epsilon} |H^{(m-j)}(x) \Omega^{(j)}(\epsilon, x)| \\ &\leq \left[\max_{|x| \leq 2\epsilon} \frac{|H^{(m-j)}(x)|}{\epsilon^j} \right] \cdot C_j \rightarrow 0 \end{aligned}$$

and $L(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$, and (7.1) is established.

We now define

$$E_{m,p}(\epsilon, x) = \Omega(\epsilon, x) \{e^{ix} - 1 - \Psi_{m,p}(x)\} + \Psi_{m,p}(x)$$

and have, for fixed m and θ such that $2m\theta > 1$,

$$\lim_{p \rightarrow \infty} \left[\lim_{\epsilon \rightarrow 0} T_\theta \{E_{m,p}(\epsilon, x)\} \right] = 0,$$

while for $|x| < \epsilon \pmod{2\pi}$, $E_{m,p}(\epsilon, x) = e^{ix} - 1$.

8. G -functions with assigned derivatives and small norm. We now seek to replace a G -function by a function of small norm C -locally equivalent to it. We begin by finding a function of small norm having the same derivatives as the given function at the origin.

Let $G(x)$ be an entire function which vanishes at the origin and m a positive integer. Then $G[\psi_m(px)/p]$ (where $\psi_m(x)$ is the function defined in (6.1)) has its first $2m$ derivatives at $x=0$ equal to those of $G(x)$, and the norm of $G[\psi_m(px)/p]$ approaches zero as $p \rightarrow \infty$.

9. G -functions locally of small norm. Let $G(x)$ be a G -function which is also an entire function having $G(0) = G'(0) = \dots = G^{(m)}(0) = 0$. Let

$$\Omega^*(\epsilon, x) = \begin{cases} \Omega(\epsilon, x), & 0 \leq |x| \leq \pi, \\ 0, & |x| \geq \pi. \end{cases}$$

Then if $1/m < \theta \leq 1$, we shall show that

$$(9.1) \quad \lim_{\epsilon \rightarrow 0} T_\theta \{ \Omega^*(\epsilon, x) G(x) \} = 0.$$

This statement is proved in much the same way as the corresponding statement for H -functions.

Let

$$\Omega^*(\epsilon, x) G(x) = P^*(\epsilon, x) = \int_{-\infty}^{\infty} p^*(\epsilon, \xi) e^{-i\xi x} d\xi,$$

so that

$$p^*(\epsilon, \xi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P^*(\epsilon, x) e^{i\xi x} dx.$$

Integrating by parts we have

$$p^*(\epsilon, \xi) = \frac{1}{(-i\xi)^m 2\pi} \int_{-\pi}^{\pi} P^{*(m)}(\epsilon, x) e^{i\xi x} dx \quad \text{if } \xi \neq 0.$$

Thus

$$|p^*(\epsilon, \xi)| < \frac{2}{|\xi|^m + 1} L^*(\epsilon) \quad \text{for all } \xi,$$

where $L^*(\epsilon)$ is the greater of the upper bounds of $|P^*(\epsilon, x)|$ and $|P^{*(m)}(\epsilon, x)|$ on $|x| \leq \pi$. Now as $\epsilon \rightarrow 0$, the bounds of $|P^*(\epsilon, x)|$ and $|P^{*(m)}(\epsilon, x)|$ approach zero for the same reason that $|P(\epsilon, x)|$ and $|P^{(m)}(\epsilon, x)|$ approached zero in §7. Thus $\lim_{\epsilon \rightarrow 0} L^*(\epsilon) = 0$, and

$$\begin{aligned} T_\theta [P^*(\epsilon, x)] &= 2 \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} |p^*(\epsilon, \xi)| d\xi \right]^\theta \\ &\leq 2 [L^*(\epsilon)]^\theta \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} \frac{2d\xi}{(1 + |\xi|^m)} \right]^\theta; \end{aligned}$$

and if $m\theta > 1$, (9.1) holds.

We now define for any G -function which is also an entire function and vanishes at the origin,

$$\Gamma_{m,p}(G | \epsilon, x) = \Omega^*(\epsilon, x) \{ G(x) - G[\psi_m(px)/p] \} + G[\psi_m(px)/p],$$

and we have for fixed m and θ such that $2m\theta > 1$,

$$\lim_{p \rightarrow \infty} \lim_{\epsilon \rightarrow 0} T_\theta \{ \Gamma_{m,p}(G | \epsilon, x) \} = 0,$$

while for $|x| < \epsilon$, $\Gamma_{m,p}(G | \epsilon, x) = G(x)$.

10. ***G*-functions of small norm at ∞ .** Turning now to the *C*-neighborhood of ∞ , we seek to replace an entire *G*-function by a *G*-function of small norm equal to the given function near ∞ . Let $G(x)$ be an entire *G*-function given by

$$G(x) = \int_{-A}^A e^{-iux} g(u) du$$

(where $g(u) = 0$ if $|u| > A$), and let

$$G_\delta(x) = \int_{-\infty}^{\infty} e^{-iux} g_\delta(u) du$$

where

$$g_\delta(u) = \frac{1}{\delta^3} \int_u^{u+\delta} \int_{\xi_1}^{\xi_1+\delta} \int_{\xi_2}^{\xi_2+\delta} g(\xi_1) d\xi_1 d\xi_2 d\xi_3$$

is the triple smoothing of $g(u)$.

Then

$$T_\theta[G(x) - G_\delta(x)] = 2 \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} |g(u) - g_\delta(u)| du \right]^\theta;$$

so

$$\lim_{\delta \rightarrow 0} T_\theta[G(x) - G_\delta(x)] = 0.$$

Now let

$$[1 - \Omega^*(N, x)]G_\delta(x) = P_\delta(N, x) = \int_{-\infty}^{\infty} p_\delta(N, \xi) e^{-i\xi x} d\xi.$$

Since $g_\delta(u)$ has two continuous derivatives, if $m > 1/\theta$, $G_\delta^{(k)}(x) = o(1/x^2)$ and $P_\delta^{(k)}(N, x) = o(1/x^2)$ at $\pm \infty$ for $k = 0, 1, \dots, m$. But

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} P_\delta(N, x) e^{iux} dx = p_\delta(N, u),$$

and integrating by parts, we have

$$p_\delta(N, u) = \frac{1}{(-iu)^m 2\pi} \int_{-\infty}^{\infty} e^{iux} P_\delta^{(m)}(N, x) dx.$$

Thus if $L_\delta(N)$ is the greater of the upper bounds of

$$|(x^2 + 1)P_\delta(N, x)|, \quad |(x^2 + 1)P_\delta^{(m)}(N, x)|,$$

we have

$$|p_\delta(N, u)| \leq \min \left(1, \frac{1}{|u|^m} \right) \frac{\pi}{2\pi} L_\delta(N) \leq \frac{L_\delta(N)}{1 + |u|^m};$$

so that

$$T_\theta[P_\delta(N, x)] = 2 \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} |p_\delta(N, \xi)| du \right]^\theta \\ \leq 2[L_\delta(N)]^\theta \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} \frac{du}{1+|u|^m} \right]^\theta,$$

where the latter sum converges if $m\theta > 1$. But $1 - \Omega^*(N, x)$ is never numerically greater than 1 and is zero when $|x| < N$; so $\overline{\text{Bd}} P_\delta(N, x) \rightarrow 0$ as $N \rightarrow \infty$. Moreover for $N > 1$, each derivative of $1 - \Omega^*(N, x)$ is bounded in N and x , and is zero when $|x| < N$. Thus each term of Leibnitz' expansion of

$$(1+x^2) \frac{d^m}{dx^m} [(1 - \Omega^*(N, x))G_\delta(x)]$$

has its upper bounds approach zero as $N \rightarrow \infty$. Hence $L_\delta(N) \rightarrow 0$ as $N \rightarrow \infty$, and $\lim_{N \rightarrow \infty} T_\theta[P_\delta(N, x)] = 0$.

Finally, we define

$$\Gamma_{N,\delta}^*(G|x) = G(x) - G_\delta(x) + P_\delta(N, x)$$

and have

$$\lim_{\delta \rightarrow 0} \lim_{N \rightarrow \infty} T_\theta[\Gamma_{N,\delta}^*(G|x)] = 0$$

and $\Gamma_{N,\delta}^*(G|x) = G(x)$ for $x \geq 2N$.

11. Proof of Theorem I. Returning now to the proof of Theorem I, we note that the fact that $\mathcal{Y}[F(x)]$ consists of n continuous functions is obvious, as is also the same fact for $\mathcal{Y}(H(x))$, since each value of $H(x)$ belongs to R . Now the symmetric functions of the n values of $\mathcal{Y}(H(x))$ are single-valued analytic functions of $\mathcal{Y}(H(x))$ which are therefore almost periodic; and by Walther's theorem on algebraic functions of almost periodic functions, it follows that each branch of $\mathcal{Y}(H(x))$ is almost periodic.

Since the closed set R^* is contained in the open set \mathcal{R} , there is a positive number η such that every number within a distance η of R^* is in \mathcal{R} . Choose L so great that $|G(x)| < \eta/2$ when $|x| \geq L$. Then $|F(x) - H(x)| < \eta/2 + |S(x)|$, so that if once chosen on the same branch, $F(x)$ and $H(x)$ remain on the same branch of $\mathcal{Y}(z)$ when $x \geq L$ and also when $x \leq -L$. Thus given $\epsilon > 0$, we can find $\delta > 0$ such that if $x > L$ and $|F(x) - H(x)| < \delta$, then $|\mathcal{Y}(F(x))_j - \mathcal{Y}(H(x))_{p_j}| < \epsilon$, where p_j is so chosen that $[\mathcal{Y}(F(x))_j]$ and $[\mathcal{Y}(H(x))_{p_j}]$ are on the same branch for $x > L$. From this (1.1) and (1.2) readily follow. But two almost periodic functions never have their mean square difference zero, and hence p_j and p'_j are unique.

Now suppose that $p_j = p'_j$; and write simply $\mathcal{Y}(F(x))$ and $\mathcal{Y}(H(x))$ for

$[\mathcal{Y}(F(x))]_j$ and $[\mathcal{Y}(H(x))]_{p_j}$. Then whenever $|x| \geq L$, $F(x)$ and $H(x)$ are to be taken on the same branch of $\mathcal{Y}(z)$.

In the closed region $\mathcal{R}_{\eta/2}^*$ of points within $\eta/2$ of R^* , there is a minimum distance γ between the branches of $\mathcal{Y}(z)$, so that for all z in $\mathcal{R}_{\eta/2}^*$, $|\mathcal{Y}(z)_j - \mathcal{Y}(z)_k| \geq \gamma$ for all pairs of branches. Now if M is the least common module of $\mathcal{Y}(H(x))$ and $H(x)$, we can find a finite number of elements μ_1, \dots, μ_q of M and a positive number ϵ_1 such that whenever

$$|\mu_j h| < \epsilon_1 \pmod{2\pi}, \quad j = 1, \dots, q,$$

then

$$|\mathcal{Y}(H(x+h)) - \mathcal{Y}(H(x))| < \gamma/2 \quad \text{for all } x,$$

and $|H(x+h) - H(x)|$ is uniformly so small that if \mathcal{Y} were taken on the same branch for both of these values, then $|\mathcal{Y}(H(x+h)) - \mathcal{Y}(H(x))|$ would also be less than $\gamma/2$. Thus $\mathcal{Y}(H(x+h))$ and $\mathcal{Y}(H(x))$ are on the same branch for all x .

Choose $\bar{\delta}$ so small that $[4\bar{\delta} + T_\theta[S(x)]]^{1/\theta} + (2\bar{\delta})^{1/\theta} < T_\theta[S(x)]^{1/\theta} + \eta/2$. Let $H(x) = \sum_{k=1}^{\infty} h_k e^{i\lambda_k x}$, and let q' be chosen so that if $H^*(x) = \sum_{k=q'+1}^{\infty} h_k e^{i\lambda_k x}$, then $T_\theta(H^*(x)) < \bar{\delta}$. We shall show that $\mathcal{Y}(F(x))$ is locally of finite norm with respect to $\mu_1, \dots, \mu_q, \lambda_1, \dots, \lambda_{q'}$ in the $\mathcal{O}_{q+q'}C$ neighborhood of every infinite point of $\mathcal{O}_{q+q'}C$; and hence after a similar argument for finite C neighborhoods that $\mathcal{Y}[F(x)]$ is of finite norm.

Let us consider the $\mathcal{O}C$ neighborhood of $(\mu_1 \xi_1, \dots, \mu_q \xi_q, \lambda_1 \xi'_1, \dots, \lambda_{q'} \xi'_{q'}; \infty)$. Let

$$S(x) = \int_{-\infty}^{\infty} e^{-i\xi x} d\xi(\xi), \quad G(x) = \int_{-\infty}^{\infty} e^{-i\xi x} g(\xi) d\xi;$$

and let $G^*(x) = \int_{-A}^A e^{-i\xi x} g(\xi) d\xi$, where A is so great that $T_\theta[G(x) - G^*(x)] < \bar{\delta}$. Let $[\Gamma_{N', \delta'}^*[G^*|x] = G^{**}(x)]$, where N', δ' are chosen so that $T_\theta[G^{**}(x)] < \bar{\delta}$. Then $G^{**}(x) = G^*(x)$ when $|x| > 2N'$. Let

$$H^{**}(x) = \sum_{k=1}^{q'} h_k E_{m, p_k}(\epsilon^{(k)}, \lambda_k x - \lambda_k \xi_k) e^{i\lambda_k \xi_k},$$

where $m, p_k, \epsilon^{(k)}$ are so chosen that

$$T_\theta[h_k E_{m, p_k}(\epsilon^{(k)}, \lambda_k x - \lambda_k \xi_k)] < \bar{\delta}/2q'.$$

Then $T_\theta[H^{**}(x)] < \bar{\delta}$, and

$$H^{**}(x) = H(x) - H^*(x) - \sum_{k=1}^{q'} h_k e^{i\lambda_k \xi_k}$$

whenever $|\lambda_k(x - \xi_k)| < \epsilon^{(k)} \pmod{2\pi}$, ($k=1, \dots, q'$). Let

$$\epsilon = \min(\epsilon_1/2, 1/2N', 1/L, \epsilon^{(1)}, \dots, \epsilon^{(q')}),$$

and consider the ϵ -neighborhood of $(\mu_1\xi_1, \dots, \mu_{q'}\xi_{q'}, \lambda_1\xi'_1, \dots, \lambda_{q'}\xi'_{q'}; \infty)$. We shall show that the function $\mathcal{Y}(F(x))$ is locally of finite norm with respect to $\mu_1, \dots, \mu_{q'}, \lambda_1, \dots, \lambda_{q'}$ in this neighborhood. We say $x \in \mathcal{N}$ if $(x\mu_1, \dots, x\mu_{q'}, x\lambda_1, \dots, x\lambda_{q'}, x)$ is in this neighborhood.

If $x \in \mathcal{N}$, we have $G^{**}(x) = G^*(x)$ and

$$(11.1) \quad H^{**}(x) = H(x) - H^*(x) - \tau$$

where $\tau = \sum_{k=1}^{q'} h_k e^{i\lambda_k \xi_k}$; and if x_1, x_2 are any two points in \mathcal{N} ,

$$|(x_1 - x_2)\mu_j| < \epsilon_1 \pmod{2\pi}, \quad j = 1, \dots, q,$$

and

$$|\mathcal{Y}(H(x_1)) - \mathcal{Y}(H(x_2))| < \gamma/2,$$

and $\mathcal{Y}[H(x_1)]$ and $\mathcal{Y}[H(x_2)]$ are on the same branch of $\mathcal{Y}(z)$. Thus $\mathcal{Y}(z)$ is a single-valued analytic function over the set of values of $H(x)$ for x in \mathcal{N} , and since there is no branch point or singularity within $[T_\theta(S(x))]^{1/\theta} + \eta/2$ of these points, it follows that for x in \mathcal{N} , $\mathcal{Y}(z)$ is analytic and single-valued for all $z = F(x)$. Moreover $\mathcal{Y}(z)$ has an analytic and single-valued branch for all points within $\eta/2 + [T_\theta(S(x))]^{1/\theta}$ of points for which $z = F(x)$ with x in \mathcal{N} , and on this branch $\mathcal{Y}[F(x)]$ has the values we have agreed to denote by $\mathcal{Y}(F(x))$.

For x in \mathcal{N} , $F(x) = F^*(x)$, where

$$F^*(x) = \tau + H^*(x) + H^{**}(x) + [G(x) - G^*(x)] + G^{**}(x) + S(x).$$

Moreover

$$(11.2) \quad T_\theta\{F^*(x) - \tau\} < 4\bar{\delta} + T_\theta[S(x)],$$

and by (11.1) the point τ is within $(2\bar{\delta})^{1/\theta}$ of R ; and by (11.2) and the definition of $\bar{\delta}$, we find that $\mathcal{Y}(z)$ is analytic and single-valued throughout a circle of radius greater than $\{T_\theta[F^*(x) - \tau]\}^{1/\theta}$ about τ . Thus by Lemma 1, $\mathcal{Y}[F^*(x)] \in A^*$, and since $\mathcal{Y}[F(x)] = \mathcal{Y}[F^*(x)]$ in \mathcal{N} , $\mathcal{Y}[F(x)]$ is locally of finite norm with respect to $(\mu_1, \dots, \mu_{q'}, \lambda_1, \dots, \lambda_{q'})$ in a $C\bar{\mathcal{O}}$ neighborhood of $(\mu_1\xi_1, \dots, \mu_{q'}\xi_{q'}, \lambda_1\xi'_1, \dots, \lambda_{q'}\xi'_{q'}; \infty)$. The simpler fact that $\mathcal{Y}[F(x)]$ is locally of finite norm in the neighborhood of every finite point of C can be proved in a similar but simpler manner; and we therefore conclude that $\mathcal{Y}[F(x)] \in A_\theta$.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
CAMBRIDGE, MASS.

NETS AND GROUPS*

BY

REINHOLD BAER

The combinatorial properties, underlying the configuration of three pencils of parallel straight lines in the plane, have found their condensation in the concept of "net." The theory of nets[†] culminates in two extreme results: Bol's theorem that every net may be represented by means of coordinates which are taken out of certain abstract multiplicative manifolds—these need not be associative—and Thomsen's characterization of those nets whose coordinates may actually be chosen from a group, which theorem started the whole theory.

The principal object of this paper is to show that the theory of nets is completely equivalent to a well-determined chapter in the theory of groups, using this term in the customary sense of the word. To do this we have to investigate certain groups of net transformations. These groups contain all the possible systems of net coordinates and provide us therefore with the means to characterize those systems of coordinates which define isomorphic nets—a net may be describable by several non-isomorphic systems of coordinates. This method leads incidentally to a rather simple proof of Thomsen's theorem and to some new characterizations of the group-nets.

The net-theoretical considerations are preceded by a systematic discussion of those multiplicative manifolds which may be derived from the multiplication of cosets in a group.[‡] Their importance for the theory of nets arises from

* Presented to the Society, November 25, 1938; received by the editors September 10, 1938.

† The following papers are concerned with the theory of nets: W. Blaschke and G. Bol, *Geometrie der Gewebe: Topologische Fragen der Differentialgeometrie*, Berlin, 1938; G. Bol, *Mathematische Annalen*, vol. 114 (1937), pp. 414–431; H. Kneser, *Abhandlungen aus dem Mathematischen Seminar, Hamburg*, vol. 9 (1932), pp. 147–151; R. Moufang, *Mathematische Annalen*, vol. 110 (1934), pp. 416–430; K. Reidemeister, *Mathematische Zeitschrift*, vol. 29 (1929), p. 427; K. Reidemeister, *Grundlagen der Geometrie*, Berlin and Leipzig, 1930; G. Thomsen, *Abhandlungen aus dem Mathematischen Seminar, Hamburg*, vol. 7 (1929), pp. 99–106. It should be noted that the nets are sometimes called "webs" (in German, "Gewebe").

‡ Generalizations of the group concept which have some bearing on our investigations have been discussed in the following papers: R. Baer, *Sitzungsberichte der Heidelberger Akademie, mathematisch-naturwissenschaftliche Klasse*, (4), 1928. G. Bol, *Mathematische Annalen*, vol. 114 (1937), pp. 414–431; H. Brandt, *Mathematische Annalen*, vol. 96 (1927), pp. 360–366; M. Dresher and O. Ore, *American Journal of Mathematics*, vol. 60 (1938), pp. 705–733; L. W. Griffiths, *American Journal of Mathematics*, vol. 60 (1938), pp. 345–354; A. Loewy, *Journal für die reine und angewandte Mathematik*, vol. 157 (1927), pp. 239–254; F. Marty, *Comptes Rendus de l'Académie des Sciences*, vol. 201 (1935), pp. 636–638; F. Marty, *Annales de l'École Normale Supérieure*, vol. 53 (1936), pp.

the fact that all the admissible systems of net coordinates are of this type.

1. **Coset multiplication.** A multiplication in the set M of elements is a single-valued* function of the ordered pairs of elements in M with values in M . If a multiplication xy has been defined for the elements of M , then M shall be called a multiplication system (with regard to this multiplication xy).

If M is a multiplication system (with regard to the multiplication xy), then a *left unit* is an element e which satisfies $ex = x$ for every element x in M . Right units are defined accordingly and elements which are right and left units at the same time are called *units*.

The multiplication system M is said to be a *left-division system*, if there exists corresponding to any pair u, v of elements in M one and only one element x in M so that $xu = v$. Right-division systems are defined accordingly and systems which are at the same time right- and left-division systems are called *division systems*.

If u is an element in the multiplication system M , then the *right translation* of M corresponding to the element u maps the element x of M upon the element xu of M . The right translations of M are one-one mappings of M upon the whole set M if, and only if, M is a left-division system, and in this case as permutations of M they generate a subgroup of the group of permutations of M .

It is our object in this section to investigate the multiplications of cosets. A fairly general type of coset multiplication may be described in the following fashion. Let S be a subgroup of the group G , and let $r(X)$ be a fixed system of representatives of the right cosets $X = Sr(X)$ of G modulo S (so that $r(X) = r(Y)$ if, and only if, $Sr(X) = Sr(Y)$). Then the multiplication system $(S < G; r(X))$ consists of the right cosets X of G modulo S , and the multiplication in $(S < G; r(X))$ is defined by the following rule:

$$XY = Sr(X)r(Y).$$

(1.0) If G' is the subgroup of G which is generated by the elements $r(X)$, and if S' is the crosscut of G' and S , then $(S < G; r(X))$ and $(S' < G'; r(X))$ are isomorphic, since every coset $Sr(X)$ contains one and only one coset of G' modulo S' (namely $S'r(X)$).

83-123; O. Ore, *Duke Mathematical Journal*, vol. 3 (1937), pp. 149-174; F. K. Schmidt, *Sitzungsberichte der Heidelberger Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Klasse*, (8), 1927, pp. 91-103; Erich Schönhardt, *Über lateinische Quadrate und Unionen*, *Journal für die reine und angewandte Mathematik*, vol. 163 (1930), pp. 183-230; H. S. Wall, *American Journal of Mathematics*, vol. 59 (1937), pp. 77-98.

* That it is no loss in generality to restrict one's attention to single-valued functions, has been pointed out by L. W. Griffiths (*American Journal of Mathematics*, vol. 60 (1938), pp. 345-354). For the induced multiplication in the set of subsets is certainly single-valued.

THEOREM 1.1. (a) *The multiplication system M is isomorphic with a system $(S < G; r(X))$ if, and only if, M is a left-division system possessing a left unit.*

(b) *The right translations of the multiplication system $M = (S < G; r(X))$ generate a group $T(M)$ of permutations of M .*

(c) *If G' is the subgroup of G , generated by the elements $r(X)$, and if S' is the crosscut of G' and S , then there exists a homomorphism κ of G' upon $T(M)$ with the following properties:*

(i) *S'^* consists of those elements in $T(M)$ which leave the left unit in M invariant.*

(ii) *The elements, mapped by κ upon the identity, form the greatest normal subgroup of G' which is a subgroup of S' .*

(iii) *κ maps the set of elements $r(X)$ upon the (whole) set of the right translations of M , and, in particular, the element $r(X)$ upon the right translation of M , corresponding to X .*

Proof. Let us consider first a multiplication system $M = (S < G; r(X))$. Then $SX = Sr(S)r(X) = Sr(X)$ for every X in M , and S is consequently a left unit in M . If U and V are two elements in M , then the solutions of the equation $XU = V$ are exactly the solutions of the equation $Sr(X)r(U) = Sr(V)$, and the solutions of this equation are the same as the solutions of the equation $Sr(X) = Sr(V)r(U)^{-1}$. Since this last equation has one and only one solution, namely $X = Sr(V)r(U)^{-1}$, it follows that M is a left-division system. This proves (b) and the necessity of the conditions in (a).

If t is any element in G' , then the right translation of G' corresponding to t induces a uniquely determined permutation t^* of the elements in $M = (S' < G'; r(X))$. (Note that $(S < G; r(X))$ and $(S' < G'; r(X))$ are essentially the same.) Since $r(X)^*$ is in particular the right translation of M corresponding to X , it follows that κ is a homomorphism of G' upon the whole group $T(M)$ which satisfies (iii). If t is any element in G' , then $S't = S'$ if, and only if, t is an element in S' , and this proves that κ satisfies (i). If E is the subgroup of G' which consists of the elements mapped by κ upon the identity, then E is a normal subgroup of G' and it follows from (i) that $E \leq S'$. If, conversely, F is a normal subgroup of G' , and if $F \leq S'$, then

$$S'r(X)f = S'r(X)fr(X)^{-1}r(X) = S'r(X)$$

for every f in F and every X in M . Hence $F^* = 1$, and this completes the proof of (ii) and of (c).

Suppose now that M is a left-division system, possessing a left unit e . Denote by $t(x)$ the right translation of M , corresponding to the element x in M , and let $T(M)$ be the group generated by the $t(x)$, and $S(M)$ the sub-

group consisting of all those elements in $T(M)$ which leave e invariant. Two elements in $T(M)$ belong to the same right coset of $T(M)$ modulo $S(M)$ if, and only if, they map e upon the same element x of M . Since there exists one and only one right translation of M which maps e upon x , namely $t(x)$, it follows that the $t(x)$ form a complete set of representatives of the right cosets of $T(M)$ modulo $S(M)$. A one-one correspondence between M and $(S(M) < T(M); t(x))$ is therefore defined in mapping the element x in M upon the element $S(M)t(x)$. This correspondence is an isomorphism, since the transformation

$$S(M)t(x)S(M)t(y) = S(M)t(x)t(y) = S(M)t(xy)$$

maps e upon xy . This completes the proof of (a), and it shows, moreover, that the following statement is true:

COROLLARY 1.2. *If M is a left-division system, possessing a left unit e , if $T(M)$ is the group generated by the right translations $t(x)$ of M , and if $S(M)$ consists of those permutations in $T(M)$ which leave e invariant, then the right translations form a complete set of representatives of the right cosets of $T(M)$ modulo $S(M)$ and an isomorphism of M upon $(S(M) < T(M); t(x))$ is defined by mapping x upon $S(M)t(x)$.*

The following statement is a simple consequence of Theorem 1.1:

COROLLARY 1.3. *An isomorphism of the group G upon $T(M) = T[(S < G; r(X))]$ is defined by mapping the element x of the group G upon the permutation x^* of the multiplication system $M = (S < G; r(X))$ which the right translation, corresponding to x , induces in M if, and only if,*

- (1) G is generated by the elements $r(X)$;
- (2) the crosscut of all the subgroups of G which are conjugate to S in G is 1.

The following statement serves to analyze the relation between the two conditions involved in Theorem 1.1 (a).

(1.4) *The left-division system M possesses a left unit if, and only if, there exist in M elements w which satisfy*

- (i) $w(xy) = (wx)y$ for all x and y in M ;
- (ii) $wx = wy$ implies $x = y$.

Proof. The condition is necessary, since the left unit satisfies (i) and (ii).

If conversely w is an element in M which satisfies (i) and (ii), then there exists one and only one solution e of $ew = w$ in M . This element e satisfies $ww = w(ew) = (we)w$ and hence $w = we$, since M is a left-division system. Furthermore $wx = (we)x = w(ex)$ and therefore $x = ex$ by (ii) for every x in M , and this proves that e is a left unit.

The coset multiplication in a system $(S < G; r(X))$ is determined by the choice of the representatives $r(X)$. That to some degree the choice of the representatives is determined by the coset multiplication may be seen from the following statement:

(1.5) *The two sets of representatives $r(X)$ and $r'(X)$ of the right cosets X of the group G modulo its subgroup S define the same multiplication of the cosets, that is, $(S < G; r(X)) = (S < G; r'(X))$ if, and only if, each of the quotients $r'(X)r(X)^{-1}$ is contained in a normal subgroup of G which is a subgroup of S .*

REMARK. *If the subgroup S of G has the property that 1 is the only normal subgroup of G which is contained in S , then the two sets $r(X), r'(X)$ of representatives define the same coset multiplication if, and only if, $r(X) = r'(X)$ for every X .*

Proof. Since $r(X)$ and $r'(X)$ are both elements in the right coset X , we have $r'(X) = s(X)r(X)$ where $s(X)$ is a suitable element in S . If the two sets of representatives define the same coset multiplication, then

$$Sr'(X)Sr'(Y) = Sr'(X)r'(Y) = Sr(X)s(Y)r(Y) = Sr(X)r(Y)$$

and consequently $Sr(X)s(Y) = Sr(X)$ or $Sr(X)s(Y)r(X)^{-1} = S$ for every pair X, Y . If now U is some right coset, g any element in G , then $g = sr(Sg)$ for some s in S and

$$S = Sr(Sg)s(U)r(Sg)^{-1} = Ssr(Sg)s(U)r(Sg)^{-1}s^{-1} = Sgs(U)g^{-1}.$$

This shows that every $gs(U)g^{-1} = gr'(U)r(U)^{-1}g^{-1}$ is contained in S , proving the necessity of our condition.

If the condition is satisfied, then

$$\begin{aligned} Sr'(X)Sr'(Y) &= Sr'(X)r'(Y) = Sr(X)s(Y)r(Y) \\ &= Sr(X)s(Y)r(X)^{-1}r(X)r(Y) \\ &= Sr(X)r(Y) = Sr(X)Sr(Y), \end{aligned}$$

and this completes the proof.

2. Division systems. The only multiplication systems we shall need for our applications are the division systems with unit. These are certainly left-division systems with left units, and they are therefore of the form $(S < G; r(X))$.

THEOREM 2.1. *The multiplication system $M = (S < G; r(X))$ possesses a unit if, and only if, all the conjugates in G to the element $r(S)$ are contained in S ; that is, if, and only if, all the elements $r(X)r(S)r(X)^{-1}$ are in S .*

Proof. If all the conjugates of $r(S)$ are in S , then

$$XS = Sr(X)r(S) = Sr(X)r(S)r(X)^{-1}r(X) = Sr(X) = X,$$

and M possesses therefore the unit S . If conversely M possesses a unit, then S is this unit. If x is any element in G , then there exists an element s in S so that $x = sr(Sx)$ and

$$Sr(Sx) = Sr(Sx)r(S) = Ssr(Sx)r(S) = Sxr(S)x^{-1}sr(Sx).$$

But this implies that $xr(S)x^{-1}s$ and consequently $xr(S)x^{-1}$ are elements in S .

REMARK 2.2. If, as we may assume without loss in generality (cf. (1.0)), the only normal subgroup of G , contained in S , is 1, then the existence of a unit in $(S < G; r(X))$ is equivalent to the fact that $r(S) = 1$.

THEOREM 2.3. The multiplication system $(S < G; r(X)) = M$ is a division system if, and only if, the elements $r(X)$ form a complete set of representatives for the right cosets of the group G modulo every subgroup of G which is conjugate to S in G .

Proof. Assume first that M is a division system. If g is any element in G , then $g = sr(Sg)$ for a suitable element s in S . If w is another element in G , then there exists one and only one element X in M so that $(Sg)X = S(gw)$, and this X is clearly the only solution of

$$(g^{-1}Sg)w = r(Sg)^{-1}Sr(Sg)w = r(Sg)^{-1}Sr(Sg)r(X) = g^{-1}Sgr(X).$$

Thus the elements $r(X)$ form a complete set of representatives for the right cosets of G modulo $g^{-1}Sg$, if M is a division system.

Suppose now conversely that the elements $r(X)$ form a complete set of representatives of the right cosets of G modulo every $g^{-1}Sg$. If U and V are two elements of M , then the solutions X of $UX = V$ are exactly the solutions X of the equation $Sr(U)r(X) = Sr(V)$ and these are exactly the solutions of

$$r(U)^{-1}Sr(U)r(X) = r(U)^{-1}Sr(U)r(U)^{-1}r(V);$$

that is, $r(X)$ is the uniquely determined representative of the right coset $r(U)^{-1}Sr(U)r(U)^{-1}r(V)$ of G modulo $r(U)^{-1}Sr(U)$. This shows that M is a right-division system and consequently a division system.

REMARK 2.4. If $(S < G; r(X))$ is a division system, and if g is any element in G , then the equation

$$U = SgX = Sr(Sg)r(X) = Sgr(X)$$

has one and only one solution X and the elements $gr(X)$ for X in $(S < G; r(X))$ form therefore a complete set of representatives of the right cosets for every fixed element g .

THEOREM 2.5. If the elements $r(X)$ form a complete set of representatives of the right cosets of the group G modulo its subgroup S , if G' is generated by the

elements $r(X)$ and S' is the crosscut of G' and S , then the following three assertions are equivalent:

- (a) S' is a normal subgroup of G' .
- (b) $(S < G; r(X))$ is a group.
- (c) $(S < G; r(X))$ is associative.*

Proof. (b) is a consequence of (a), since $(S < G; r(X))$ and $(S' < G'; r(X))$ are isomorphic. (c) is obviously a consequence of (b). Assume finally that $(S < G; r(X))$ is associative. Then

$$Sr(Z)r(X)r(Y) = [Sr(Z)r(X)]Y = (ZX)Y = Z(XY) = Sr(Z)r(XY),$$

and $r(Z)r(XY)r(Y)^{-1}r(X)^{-1}r(Z)^{-1}$ is therefore, for every triple X, Y, Z , an element in S' . If N is the greatest normal subgroup of G' which is contained in S' , then it follows from this remark that $r(XY)r(Y)^{-1}r(X)^{-1}$ is contained in N and that consequently $Nr(XY) = Nr(X)Nr(Y)$. The classes $Nr(X)$ form, therefore, a subset of the group G'/N which is closed with regard to multiplication. Since $(S' < G'; r(X))$ is a left-division system, there exists for every X one and only one X^{-1} so that $r(S) = r(X^{-1}X)$. Since $(S' < G'; r(X))$ is an associative left-division system with left unit S' , we have $XX = X(S'X) = (XS')X$ and therefore $X = XS'$; that is, S' is the unit of the system and $r(S)$ is therefore, by Theorem 2.1, an element of N . Hence $N = Nr(X^{-1})r(X)$ or $Nr(X^{-1}) = Nr(X)^{-1}$. Consequently $Nr(X)Nr(X^{-1}) = N$. This shows that the elements $Nr(X)$ form a subgroup of G'/N . Since G' is generated by the elements $r(X)$, the elements $Nr(X)$ form the complete group G'/N . Since these elements $Nr(X)$ form a set of representatives of the right cosets of G'/N modulo S'/N , this proves that $S' = N$ is a normal subgroup of G' and thus (a) is a consequence of (c).

The following example of a division system without unit is of interest because of Theorem 6.1. The elements of the system are u, v , and w , and the multiplication table is

$$uv = vu = w^2 = w, \quad vw = wv = u^2 = u, \quad wu = uw = v^2 = v.$$

3. Similar division systems. For future application we need an extension of the concept of isomorphism. The following statements form a basis for this concept of similarity.

* It is a consequence of Theorem 1.1 (a) that the inference (c)-(b) may be stated in the following form: *An associative left-division system with left unit is a group.* A direct proof of this fact may be indicated: If e is the left unit and x any element, then $xe = xe^2 = (xe)e$; that is, $x = xe$ and e is the unit. If x is any element and x^{-1} is the uniquely determined element so that $x^{-1}x = e$, then $x = xe = x(x^{-1}x) = (xx^{-1})x = ex$ and therefore $xx^{-1} = e$; thus x^{-1} is the inverse of x and the system is a group.

(3.1) If $(S < G; r(X))$ is a division system with unit, then each

$$(g^{-1}Sg < G; r(U)^{-1}r(X)),$$

for fixed U and variable X , is a division system with unit.

Proof. It is a consequence of Theorem 2.1 and Theorem 2.3 that

$$(g^{-1}Sg < G; r(X))$$

is a division system with unit. It is a consequence of Remark 2.3 that the elements $r(U)^{-1}r(X)$ for fixed U and variable X form a complete set of representatives for the right cosets of G modulo $g^{-1}Sg$. Since $1 = r(U)^{-1}r(U)$, it follows, therefore, from Theorem 2.2 that $(g^{-1}Sg < G; r(U)^{-1}r(X))$ is a division system with unit.

If 1 is the only normal subgroup of G which is contained in the subgroup S of G , and if G is generated by the set of representatives $r(X)$ of the right cosets X of G modulo S , then G , S , and $r(X)$ are said to define a *canonical representation* of the multiplication system $M = (S < G; r(X))$. It is a consequence of Corollary 1.3 that any two canonical representations of M are isomorphic,* and it is a consequence of Corollary 1.2 and Theorem 1.1 (a) that the multiplication system M possesses a canonical representation if, and only if, M is a left-division system with left unit.

(3.2) If M is a division system with unit, if $(S < G; r(X))$ is a canonical representation of M , then $g^{-1}Sg$, G , and the elements $r(U)^{-1}r(X)$ define a canonical representation of $(g^{-1}Sg < G; r(U)^{-1}r(X))$.

Proof. As M possesses a unit and as the only normal subgroup of G which is contained in S is 1, it follows from Theorem 2.1 that $r(S) = 1$. Hence $r(U)^{-1}$ is one of the elements $r(U)^{-1}r(X)$, and these elements generate, therefore, the same group as the $r(X)$.

DEFINITION 3.3. The division system M with unit and the multiplication system M' are similar if M' is isomorphic with $(g^{-1}Sg < G; r(U)^{-1}r(X))$ where $(S < G; r(X))$ is a canonical representation of M .

If M is a division system with unit, and if the multiplication system M' is similar to M , then it follows from (3.1) that M' is a division system with unit. If furthermore M' is isomorphic with

$$(g^{-1}Sg < G; r(U)^{-1}r(X)),$$

and $(S < G; r(X))$ is a canonical representation of M , it follows from (3.2) that

$$(g^{-1}Sg < G; r(U)^{-1}r(X))$$

* Two representations $(S < G; r(X))$ and $(T < H; h(X))$ of the same system M are isomorphic if there exists an isomorphism κ of the group G upon the group H so that $S^\kappa = T$ and $r(X)^\kappa = h(X)$.

is a canonical representation of M' . This implies that the similarity relation is symmetric. That it is transitive follows from

$$(\tau(U)^{-1}\tau(V))^{-1}\tau(U)^{-1}\tau(X) = \tau(V)^{-1}\tau(X).$$

That it finally is preserved under isomorphisms follows from the fact that any two canonical representations of a multiplication system are isomorphic.

In the following fashion one will be led to another characterization of the similar division systems with unit, a characterization which is of a more intrinsic type than the one given above. If M is a division system with unit and $M = (S < G; \tau(X))$ is a canonical representation of M , then all the subgroups conjugate to S may be represented in the form $\tau(V)^{-1}S\tau(V)$. Since transformation with $\tau(V)$ induces an automorphism of G , it follows that all the similar multiplication systems may be represented (in canonical form) in the following fashion:

$$M' = (S < G; \tau(V)\tau(U)^{-1}\tau(X)\tau(V)^{-1}).$$

Denote now by X/V the uniquely determined solution Y of the equation $YV = X$ and by $X^{[V, U]}$ the uniquely determined solution Z of the equation $XV = (V/U)Z$; then the right coset of G modulo S which is represented by

$$\tau(V)\tau(U)^{-1}\tau(X)\tau(V)^{-1}$$

is just the element $((V/U)X)/V$ of M . The multiplication used so far is the multiplication in M . If X' and Y' are two elements in M' , then there exist elements X and Y so that $X' = ((V/U)X)/V$ and $Y' = ((V/U)Y)/V$, namely $X = X'^{[V, U]}$ and $Y = Y'^{[V, U]}$, and the M' -product of the elements X' and Y' is represented in G by the element

$$\tau(V)\tau(U)^{-1}\tau(X)\tau(U)^{-1}\tau(Y)\tau(V)^{-1}.$$

The M' -product of X' and Y' is therefore

$$X' \cdot Y' = \frac{\frac{X'V}{U} Y'^{[V, U]}}{V},$$

and this shows that one may get all the division systems with unit which are similar to M by choosing two elements U and V in M quite at random and then defining a new multiplication by the above formula.

The two special cases $U=1$ and $V=1$ may be stated. If $U=1$, then $X' \cdot Y' = (X'V)Y'^{[V, U]}/V$ where $Y'^{[V, U]}$ is the uniquely determined solution of the equation $Y'V = VY'^{[V, U]}$. If $V=1$ then $X' \cdot Y' = (X'/U)Y'^{[1, U]}$, where $Y'^{[1, U]}$ is the uniquely determined solution of the equation $Y' = (1/U)Y'^{[1, U]}$.

The two following remarks concern important special cases of classes of similar division systems with unit. If the two division systems M and M' , each containing a unit, are similar, and if M is a *group*, then it follows from Theorem 2.4 that M and M' are *isomorphic* groups. Suppose now that M is a division system with unit and that every system M' which is similar to M is *commutative*. Then we represent M in the canonical form $(S < G; r(X))$. As M is a division system with unit, it follows that the $r(X)$ form a complete set of representatives of the cosets modulo gSg^{-1} for any g in G . From our hypothesis, $(gSg^{-1} < G; r(X))$ is commutative. Hence

$$r(X)r(Y)r(X)^{-1}r(Y)^{-1}$$

is an element in gSg^{-1} . But as $S, G, r(X)$ is a canonical representation of M , it follows that 1 is the only element contained in every gSg^{-1} , and consequently we have

$$1 = r(X)r(Y)r(X)^{-1}r(Y)^{-1}.$$

Since G is generated by the elements $r(Z)$, this implies that G is a commutative group, and this implies that M is a *commutative group* (so that all the systems, similar to M , are isomorphic to M).

Our treatment of the left-division systems with left unit (§1) was essentially nothing else than a generalization of the proof of Cayley's theorem that every group may be represented as an isomorphic group of permutations. For this proof one uses the so-called regular representation of the group which consists just of the right translations. One is led to another generalization of this idea by restricting one's attention to the system of the right translations and not extending this system, as has been done in §1, to the generated group of permutations. As this will give us some better insight into the concept of similarity, it will be useful to consider this in some detail.

A set P of permutations of the (finite or infinite) set T of elements shall be called *simply transitive* if there exists to every pair of elements in T one and only one permutation in P which maps the one element upon the other.

The right translations of the multiplication system M form a set $P(M)$ of permutations of M if, and only if, M is a left-division system. $P(M)$ is simply transitive if, and only if, M is a division system; and $P(M)$ contains the identity if, and only if, M possesses a right unit.

The set P of permutations of the set T and the set P' of permutations of the set T' are said to be *similar* if there exists a one-one correspondence p which maps P upon P' and a pair of one-one correspondences t, s which both map T upon T' so that $tx^p = xs$ for every x in P . If in particular $s = t$, then the systems are *isomorphic*.

THEOREM 3.4. *The set P of permutations of the set T is isomorphic to the set $P(M)$ of right translations of a suitable division system M with unit if, and only if, P is simply transitive and contains the identity.*

Proof. The necessity of the conditions has been pointed out before. If the conditions are satisfied, then choose an element e in T and denote by M the system which consists of the permutations in P , where the multiplication in M has been defined in the following fashion: If x and y are two permutations in P , then $x * y$ is the uniquely determined element in P which maps e upon $e^{xy} = (e^x)^y$. This product definition in M is certainly unique. The identity in P gives rise to the unit in M . If u and v are two elements in M , then the solution of $u * x = v$ is just the uniquely determined permutation in P which maps e^u upon e^v , and M is therefore a right-division system. There exists, furthermore, a uniquely determined element f in T so that $f^u = e^v$, and there exists a uniquely determined element x in P so that $e^x = f$. Clearly this permutation x is the uniquely determined solution of the equation $x * u = v$, and M is consequently a division system with unit.

We define t by the equation $x^t = e^x$ for x in M . Since the elements x in M are the permutations of a simply transitive system, t is a one-one correspondence mapping M upon T . Furthermore, let us denote by p the correspondence which maps the right translation of M which is induced by the element u of M upon the element u (of M and) of P . The correspondence p is a one-one correspondence mapping $P(M)$ upon P , since M is a division system with unit. If now x is an element in M , u an element in $P(M)$, then

$$x^{tu^p} = (e^x)^{u^p} = e^{xu^p} = (xu^p)^t = (x^u)^t = x^{ut},$$

and this proves that t and p together define an isomorphism of $P(M)$ upon P . That proves our theorem.

THEOREM 3.5. *The two division systems M and M' , both of them possessing a unit, are similar if, and only if, $P(M)$ and $P(M')$ are similar systems of permutations.*

Proof. As both M and M' are division systems with unit, there exist canonical representations $M = (S < G; r(X))$ and $M' = (S' < G'; r'(X))$ of these systems. The system $P(M)$ is by (1.5) exactly the system of permutations which the right translations, induced by the elements $r(X)$, induce in the cosets of G modulo S ; and $P(M')$ may be described accordingly. If M and M' are similar, we may assume without loss in generality that M' is of the form

$$M' = (S < G; r(V)r(U)^{-1}r(X)r(V)^{-1}),$$

as the inner automorphism $g \rightarrow r(V)gr(V)^{-1}$ of G maps

$$(r(V)^{-1}Sr(V) < G; r(U)^{-1}r(X))$$

exactly upon

$$(S < G; r(V)r(U)^{-1}r(X)r(V)^{-1}).$$

The correspondences s and t are defined by the formulas

$$[Sr(V)r(U)^{-1}r(X)r(V)^{-1}]^s = Sr(X) \quad (= X),$$

$$[Sr(V)r(U)^{-1}r(X)r(V)^{-1}]^t = Sr(X)r(U)^{-1}.$$

Both s and t are one-one correspondences which map M' upon the whole M . If

$$Z' = Sr(V)r(U)^{-1}r(Z)r(V)^{-1}$$

is an element in M' , then the right translation induced by Z' has the form

$$[Sr(V)r(U)^{-1}r(X)r(V)^{-1}]^{Z'} = Sr(V)r(U)^{-1}r(X)r(U)^{-1}r(Z)r(V)^{-1},$$

and the correspondence p is defined as mapping this right translation of M' upon the right translation $[Sr(X)]^{Z'^p} = Sr(X)r(Z)$, or $Z'^p = Z$. The correspondence p is, by its definition and by the fact that M and M' are division systems with unit, a one-one correspondence which maps $P(M')$ upon the whole $P(M)$. Finally we have

$$\begin{aligned} [Sr(V)r(U)^{-1}r(X)r(V)^{-1}]^{Z'^s} &= Sr(X)r(U)^{-1}r(Z) \\ &= [Sr(X)r(U)^{-1}]^Z = [Sr(X)r(U)^{-1}]^{Z'^p} \\ &= [Sr(V)r(U)^{-1}r(X)r(V)^{-1}]^{tZ'^p}, \end{aligned}$$

and this shows that s , t , and p induce a similarity between $P(M')$ and $P(M)$.

Assume now conversely that s , t , and p induce a similarity between $P(M')$ and $P(M)$. If X' is an element in M' , then p maps the right translation of M' which is induced by X' upon a right translation of M which is induced by a uniquely determined element X'^p of M . Then a correspondence w may be defined as follows:

$$[S'r'(X')]^w = r(1^s)^{-1}Sr(1^s)r(1^p)^{-1}r(X'^p).$$

We put $1^s = V$ and $1^p = U$, and the above formula then reads

$$[S'r'(X')]^w = r(V)^{-1}Sr(V)r(U)^{-1}r(X'^p).$$

The correspondence w is a one-one correspondence which maps $M' = (S' < G'; r'(X'))$ upon the whole system $M'' = (r(V)^{-1}Sr(V); r(U)^{-1}r(X))$, and M'' is a division system with unit which is similar to $M = (S < G; r(X))$.

Furthermore we have $X'^t Y'^p = [X' Y']^s$, since the left-hand side signifies the effect of the right translation corresponding to Y'^p upon the element X'^t ,

and the right-hand side gives the picture under s of the effect of the right translation corresponding to Y' upon X' . Thus we have, in particular,

$$V = 1^s = 1'^p = 1'^U, \quad X'1^p = X'^U = X'^s, \quad 1'^X^p = X'^s = X'^U.$$

Hence we have

$$\begin{aligned} [S'r'(X')]^w &= r(V)^{-1}Sr(1'^U)r(U)^{-1}r(X'^p) \\ &= r(V)^{-1}Sr(1')r(X'^p) = r(V)^{-1}Sr(X'^s). \end{aligned}$$

Thus we find that

$$\begin{aligned} [X'Y']^w &= [S'r'(X'Y')]^w = r(V)^{-1}Sr([X'Y']^s) \\ &= r(V)^{-1}Sr(X'^s)r(Y'^p) \\ &= r(V)^{-1}Sr(X'^s)r(1^p)r(1^p)^{-1}r(Y'^p) \\ &= r(V)^{-1}Sr(1')r(X'^p)r(1^p)^{-1}r(Y'^p) \\ &= r(V)^{-1}Sr(1')r(1^p)r(1^p)^{-1}r(X'^p)r(1^p)^{-1}r(Y'^p) \\ &= r(V)^{-1}Sr(V)r(U)^{-1}r(X'^p)r(U)^{-1}r(Y'^p) \\ &= X'^wY'^w, \end{aligned}$$

and M' and M'' are therefore isomorphic. Hence M' and M are similar, and this completes the proof.

4. Net translations. A net consists of four different kinds of elements: points, R -lines, S -lines, and T -lines. Points may lie on lines, lines may pass through points, and lines may have points in common. These relations are subject to the following two postulates:

I. *Through every point there passes one and only one R -line, one and only one S -line, and one and only one T -line.*

II. *If the lines X and Y belong to different ones of the pencils R , S , and T , then they have one and only one point in common.*

It is a consequence of I that lines in the same pencil do not meet.

A typical example for such a net consists of the points of the plane $x+y+z=0$ in euclidean 3-space and the lines $x=\text{const.}$, $y=\text{const.}$, and $z=\text{const.}$ in this plane.

If p is a point in the net N , then the uniquely determined R -line through p is denoted by $R(p)$, and $S(p)$ and $T(p)$ are defined accordingly. If the lines X and Y belong to different ones of the pencils R , S , and T , then the uniquely determined point of the net through which both X and Y pass is denoted by $XY=YX$.

The following two formulas are recorded for future reference. Their proof is obvious.

(4.1) If X is an R -line and Y an S -line (or a T -line), then $R(XY) = X$.

(4.2) If X is an R -line and Y an S -line, then $XT(XY) = XY$.

Net isomorphisms are one-one correspondences between the points of a net N and the points of a net N' so that points on the same R -line are mapped upon points on the same R' -line, and so on. The more general kind of net isomorphism which permutes the three line pencils will not be discussed in this investigation.*

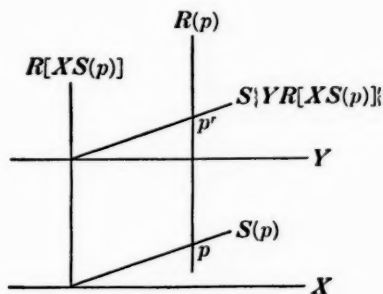


FIG. 1

Net isomorphisms, on the other hand, prove in certain respects too narrow for our purposes. Thus we introduce the following concept. A one-one correspondence t between the points of the net is termed an R/S -transformation of the net, if it satisfies the following conditions:

(a) t maps the set of all the net-points upon the whole set of all the net-points.

(b) $R(p') = R(p)$.

(c) $S(p) = S(q)$ if, and only if, $S(p') = S(q')$.

Thus R/S -transformations are characterized by the facts that they leave every R -line invariant and induce a permutation of the S -lines.† They clearly form a group, and this group is essentially the same as the group of all the permutations of the S -lines.

THEOREM 4.3. *Corresponding to every pair X, Y of T -lines there exists one and only one R/S -transformation $r(X-Y) = r(R/S; X-Y)$ which maps the points of X upon the points of Y . If p is any net-point, then p is mapped by $r(X-Y)$ upon the point $R(p)S\{YR[XS(p)]\}$.*

The theorem is illustrated by Fig. 1.

* An exception is Theorem 8.1.

† Note that nothing has been said concerning T -lines.

Proof. Assume first that t is an R/S -transformation mapping the points of the T -line X upon the points of the T -line Y . Then

$$p = R(p)S(p) = R(p)S[XS(p)] = R(p)S\{XR[XS(p)]\};$$

therefore

$$p^t = R(p)^tS\{XR[XS(p)]\}^t = R(p)S\{X^tR[XS(p)]^t\} = R(p)S\{YR[XS(p)]\}.$$

This proves that there exists at most one R/S -transformation which maps X upon Y , and that an R/S -transformation, mapping X upon Y , has the form given in the theorem.

In order to prove the existence of an R/S -transformation which maps X upon Y , let us consider, therefore, the transformation of the points of the net which is defined by

$$p^r = R(p)S\{YR[XS(p)]\}.$$

This correspondence r is certainly a single-valued function of the net-points which leaves every R -line invariant. Since

$$\begin{aligned} R(p^r)S\{XR[YS(p^r)]\} &= R(p)S\{XR[YS(R(p)S\{YR[XS(p)]\})]\} \\ &= R(p)S\{XR[YS\{YR[XS(p)]\}]\} \\ &= R(p)S\{XR\{YR[XS(p)]\}\} \\ &= R(p)S\{XR[XS(p)]\} \\ &= R(p)S[XS(p)] \\ &= R(p)S(p) = p, \end{aligned}$$

it follows that r is a one-one correspondence between the points of the net, which maps the set of the net-points upon the whole set of all the net-points. From the above formula it follows, furthermore, that

$$S(p) = S\{XR[YS(p^r)]\}, \quad S(p^r) = S\{YR[XS(p)]\},$$

and consequently $S(p) = S(q)$ if, and only if, $S(p^r) = S(q^r)$.

If finally p is a point on X , then

$$p^r = R(p)S[YR(p)] = YR(p),$$

and if p^r is a point on Y , then it follows from the above formula that

$$p = R(p^r)S[XR(p^r)] = XS(p^r).$$

This proves that r maps the points of X exactly upon the points of Y .

THEOREM 4.4. *If E is some T -line, and if u and v are points on the same R -line, then there exists one and only one R/S -transformation which maps u*

upon v and E upon some well-determined T -line. The transformation meeting the requirements is

$$r(R/S; E - T\{S(v)R[ES(u)]\}).$$

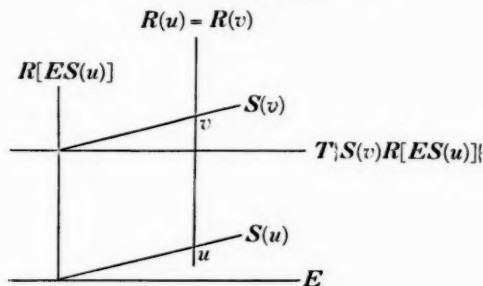


FIG. 2

Proof. If $r(E-Z)$ meets the requirements, then it follows from Theorem 4.3 that $v = R(u)S\{ZR[ES(u)]\}$. Hence

$$\begin{aligned} T\{S(v)R[ES(u)]\} &= T\{S(R(u)S\{ZR[ES(u)]\})R[ES(u)]\} \\ &= T\{S(ZR[ES(u)])R[ES(u)]\} = T\{ZR[ES(u)]\} = Z, \end{aligned}$$

and this proves the statements concerning uniqueness. Furthermore it follows from Theorem 4.3 that $r(R/S; E - T\{S(v)R[ES(u)]\})$ maps u upon the point

$$R(u)S(T\{S(v)R[ES(u)]\}R[ES(u)]) = R(v)S(v) = v,$$

as $R(u) = R(v)$, and this completes the proof.

The following statement is an obvious consequence of Theorem 4.4:

COROLLARY 4.5. *If E is a T -line and U and V are two S -lines, then there exists one and only one R/S -transformation which maps U upon V and E upon some well-determined T -line. The transformation meeting the requirements is $r(R/S; E - T\{VR[EU]\})$.*

5. Division systems and their canonical representation by net transformations. Since the transformations $r(R/S; X-Y)$ are permutations of the points in the net, they generate a group $G(R/S)$. Every element in $G(R/S)$ is a permutation of the net-points, leaves each R -line invariant, and maps every S -line upon some well-determined S -line. Concerning the T -lines not much can be said.

The transformations $r(X-Y)$ satisfy the following rules:

$$\begin{aligned} r(X-Y)r(Y-Z) &= r(X-Z), & r(Y-X) &= r(X-Y)^{-1}, \\ r(X-Y) &= r(E-X)^{-1}r(E-Y). \end{aligned}$$

The last formula implies that $G(R/S)$ is already generated by the transformations $r(E-X)$ for some fixed T -line E and variable T -lines X .

If e is some point in the net, then the transformations in $G(R/S)$ which have e as a fixed point form a subgroup $G(R/S; e)$ of $G(R/S)$.

If s and t are two elements in $G(R/S)$, then they map the point e upon the same point (of $R(e)$) if, and only if, st^{-1} is an element in $G(R/S; e)$. If p is a point on $R(e)$, then $r(R/S; T(e)-T(p))$ is the uniquely determined transformation $r(T(e)-X)$ which maps e upon p . The transformations $r(R/S; T(e)-X)$ form therefore a complete set of representatives of the right cosets of $G(R/S)$ modulo $G(R/S; e)$ and a one-one correspondence between these right cosets on the one side and the points on $R(e)$ on the other side is defined in mapping the transformations t in $G(R/S)$ with $e^t = p$ upon p .

THEOREM 5.1. (a) *The coset multiplication system*

$$M(R/S; e) = (G(R/S; e) < G(R/S; r(R/S; T(e) - X))$$

is a division system with unit, and $G(R/S)$, $G(R/S; e)$, $r(R/S; T(e)-X)$ define a canonical representation of $M(R/S; e)$.

(b) *The systems $M(R/S; p)$ form a complete set of similar division systems.*

The proof of this theorem is based on several statements some of which are of interest in themselves.

(5.1.1) *If p and q are two points in the net, then*

$$G(R/S; q) = r(R/S; T(p) - T[S(q)R(p)])^{-1}G(R/S; p) \cdot r(R/S; T(p) - T[S(q)R(p)]).$$

Since $r(R/S; T(p)-X)$ maps p upon $R(p)X$, it follows that

$$r(R/S; T(p) - X)^{-1}G(R/S; p)r(R/S; T(p) - X) = G(R/S; R(p)X).$$

Since $G(R/S; e)$ maps each R -line and $S(e)$ upon itself, it follows that every point on $S(e)$ is a fixed point under the transformations in $G(R/S; e)$ and consequently $G(R/S; e) = G(R/S; f)$ if $S(e) = S(f)$. The statement (5.1.1) is a consequence of these two special cases.

(5.1.2) *The subgroups $G(R/S; p)$ form a complete set of conjugate subgroups of $G(R/S)$.*

For if t is any element in $G(R/S)$, then, as has been remarked before, $t = t'r(R/S; T(p)-T(p'))$ for a suitable element t' in $G(R/S; p)$. Hence

$$t^{-1}G(R/S; p)t = r(R/S; T(p) - T(p'))^{-1}G(R/S; p)r(R/S; T(p) - T(p')),$$

and (5.1.2) is now a consequence of (5.1.1).

(5.1.3) *The crosscut of the groups $G(R/S; p)$ is 1; and 1 is therefore the greatest normal subgroup of $G(R/S)$ contained in $G(R/S; e)$.*

This is a consequence of (5.1.2) and the fact that a transformation t which is contained in every $G(R/S; p)$ has every net-point p as a fixed point.

Since the $r(R/S; T(e) - X)$ form a complete set of representatives of the right cosets of $G(R/S)$ modulo $G(R/S; S(p)T(e))$, as has been remarked before, and since $G(R/S; S(p)T(e)) = G(R/S; p)$ by (5.1.1), it follows from Theorem 2.3 that $M(R/S; e)$ is a division system, and it possesses a unit since $r(R/S; T(e) - T(e)) = 1$. The given representation of $M(R/S; e)$ is a canonical representation, as follows from the definition of $G(R/S)$, and a remark added to this definition, and from (5.1.3). That the multiplication sys-

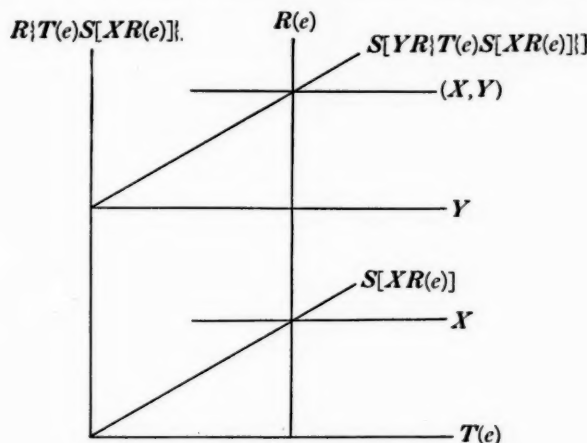


FIG. 3

tems $M(R/S; p)$ form a complete system of similar division systems is a consequence of (5.1.2) and the fact that

$$r(R/S; T(e) - U)^{-1}r(R/S; T(e) - X) = r(R/S; U - X)$$

and $G(R/S; p) = G(R/S; S(p)U)$, which completes the proof of the theorem.

The following formula will prove useful in applications:

(5.2) *If X and Y are two T -lines and (X, Y) is the T -line defined by the equation $(X, Y) = T\{R(e)S[YR\{T(e)S[R(e)X]\}]\}$, then*

$$\begin{aligned} [G(R/S; e)r(R/S; T(e) - X)][G(R/S; e)r(R/S; T(e) - Y)] \\ = G(R/S; e)r(R/S; T(e) - (X, Y)). \end{aligned}$$

This statement is illustrated by Fig. 3 above.

Proof. In order to prove this it is sufficient to show that the transformations

$$r(R/S; T(e) - X)r(R/S; T(e) - Y), \quad r(R/S; T(e) - (X, Y))$$

map e upon the same point. It follows from Theorem 4.3 that $r(R/S; T(e) - X)$ maps e upon the point

$$R(e)S\{XR[T(e)S(e)]\} = R(e)S[XR(e)] = XR(e)$$

and that therefore $r(T(e) - X)r(T(e) - Y)$ maps e upon the point

$$R(e)S\{YR\{T(e)S[XR(e)]\}\}.$$

But $r(T(e) - (X, Y))$ maps e , by Theorem 4.3, upon the point $R(e)(X, Y)$, and this proves our statement.

THEOREM 5.3. (a) *If X is a T -line and $X^R = R\{T(e)S[XR(e)]\}$, then*

$$X = T\{R(e)S[X^RT(e)]\}, \quad S[XR(e)] = S[X^RT(e)].$$

(b) *An anti-isomorphism of $M(R/S; e)$ upon $M(T/S; e)$ is defined in mapping $G(R/S; e)r(R/S; T(e) - X)$ upon $G(T/S; e)r(T/S; R(e) - X^R)$.*

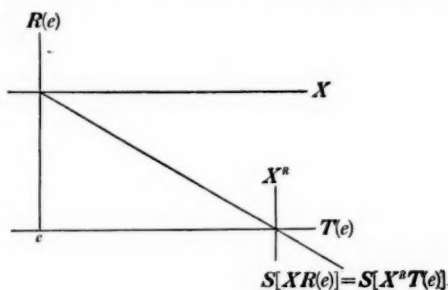


FIG. 4

Proof.* If X is a T -line, then

$$S[X^RT(e)] = S(T(e)R\{T(e)S[XR(e)]\}) = S\{T(e)S[XR(e)]\} = S[XR(e)]$$

and therefore

$$T\{R(e)S[X^RT(e)]\} = T\{R(e)S[XR(e)]\} = T[XR(e)] = X;$$

and this proves (a).

It is a consequence of (a) that the correspondence, defined in (b), is a one-one correspondence between $M(R/S; e)$ and $M(T/S; e)$. It is a conse-

* Cf. G. Bol, *op. cit.*, p. 419.

quence of (5.2) that this correspondence is an anti-isomorphism, provided, in the notation of (5.2), that $(X, Y)^R = (Y^R, X^R)$. But

$$\begin{aligned}(X, Y)^R &= R\{T(e)S[R(e)T\{R(e)S[YR\{T(e)S[R(e)X]\}\}]\}\} \\ &= R(T(e)S\{R(e)S[YR\{T(e)S[R(e)X]\}\}]\} \\ &= R\{T(e)S[YR\{T(e)S[R(e)X]\}\}\} = R\{T(e)S[YX^R]\} \\ &= R\{T(e)S[X^RT\{R(e)S[Y^RT(e)]\}\}\} = (Y^R, X^R),\end{aligned}$$

where, in order to be applicable on $M(T/S; e)$, in the formula (5.2) the symbols R and T have to be interchanged. This completely proves Theorem 5.3.

LEMMA 5.4. *If s is an R/S -transformation and t is an S/R -transformation, then $st = ts$.*

Proof. If p is any point in the net, then

$$p^{st} = (R(p)S(p))^{st} = R(p)S(p)^s = R(p)^{ts}S(p)^t = p^{ts}.$$

6. **Representation of nets.** If M is a multiplication system, then a configuration $N(M)$ may be derived from M in the following fashion. The points of $N(M)$ are the ordered pairs (x, y) of elements x, y in M . The R -lines as well as the S -lines and T -lines of the net are in one-one correspondence to the elements in M so that on the R -line corresponding to the element z in M lie the points (z, y) ; on the S -line corresponding to the element z in M lie exactly those points (x, y) which satisfy $xy = z$; and on the T -line corresponding to the element z in M lie exactly the points (x, z) .

THEOREM 6.1. *$N(M)$ is a net if, and only if, M is a division system.*

REMARK. *It is noteworthy that the existence of a unit in M is not needed here.**

Proof. The R -line, corresponding to u and the S -line corresponding to v have one and only one point in common if, and only if, the equation $ux = v$ has one and only one solution x in M ; and the T -line corresponding to u and the S -line corresponding to v have one and only one point common if, and only if, the equation $yu = v$ has one and only one solution y in M . It is obvious now how to complete the proof.†

THEOREM 6.2. (a) *The multiplication system M is a system $M(R/S; e)$ for some point e in some net N if, and only if, M is a division system with unit.*

(b) *If M is a division system with unit, if the subgroup H of the group G and the representatives $r(X)$ of the right cosets X of G modulo H form a canonical representation ($H < G; r(X)$) of M , if e is the point $(1, 1)$ of the net $N(M)$, then*

* But compare on the other hand Bol's theorem or Theorem 6.3 below.

† Cf. Bol, op. cit., p. 420.

there exists an isomorphism κ of G upon the whole group $G(R/S)$, defined for the net $N(M)$, which maps H upon $G(R/S; e)$ and $r(X)$ upon $r(R/S; T(e) - X)$, where X denotes the T -line in $N(M)$ corresponding to the element X in $M = (H < G; r(X))$; and κ induces therefore an isomorphism of M upon $M(R/S; e)$.

Proof. That the systems $M(R/S; e)$ are division systems with unit, has been proved in Theorem 5.1, and this shows that the conditions of (a) are necessary ones. Assume now that M is a division system with unit 1. Then $N(M)$ is a net by Theorem 6.1. The line $T(e)$ for $e = (1, 1)$ corresponds to the unit 1 in M . The transformation $r(R/S; T(e) - X)$, where X is the T -line corresponding to the element x in M , maps the point $p = (u, v)$ by Theorem 4.3 upon the point $R(p)S\{XR[T(e)S(p)]\}$. But $T(e)S(p) = (uv, 1)$ and therefore $XR[T(e)S(p)] = (uv, x)$, so that finally

$$R(p)S\{XR[T(e)S(p)]\} = (u, f(u, v; x))$$

where $f(u, v; x)$ is the uniquely determined solution f of the equation $uf = (uv)x$. This shows in particular that the point $(1, v)$ is mapped by $r(R/S; T(e) - X)$ upon the point $(1, vx)$.

Since $G, H, r(X)$ form a canonical representation of M , it follows from Corollary 1.3 that we may assume that

- (a) $r(X)$ is the right translation of M mapping the element v in M upon the element vx ,
- (b) G is the group of permutations of M which is generated by the right translations of M , and
- (c) H consists of those permutations in G which leave 1 invariant.

Then the right coset X of G modulo H consists of exactly those elements in G which map the unit 1 upon the element x , and the elements in the coset product XY map 1 upon xy . Thus it follows that an isomorphism of M upon $M(R/S; e)$ is defined in mapping the element x in M upon the right coset X of $G(R/S)$ modulo $G(R/S; e)$ whose elements map the point $e = (1, 1)$ upon the point $(1, x)$.

The statements of (b) are a consequence of Theorem 5.1 (a), of Corollary 1.3, and of (1.5). (a) is now a consequence of (b).

THEOREM 6.3.* *An isomorphism of the net N upon the net $N[M(R/S; e)]$ is defined in mapping the point p of N upon the point*

$$(G(R/S; e)r(R/S; T(e) - T\{R(e)S[T(e)R(p)]\}), G(R/S; e)r(R/S; T(e) - T(p)))$$

of the net $N[M(R/S; e)]$, and this isomorphism maps the point e upon $(1, 1)$.

* Cf. Bol, op. cit., pp. 418-419.

Proof. Denote by $(x(p), y(p))$ the image of the point p in N under the transformation, defined in the theorem. Then it is obvious that the points p and q lie on the same T -line if, and only if, $y(p) = y(q)$. It follows from Theorem 5.3 (a) that

$$R(p) = R\{T(e)S[R(e)R(p)^T]\}$$

if

$$R(p)^T = T\{R(e)S[T(e)R(p)]\};$$

and consequently p and q are on the same R -line if, and only if, $x(p) = x(q)$.

Since

$$\begin{aligned} T(R(e)S\{T(p)R[T(e)S(R(e)T\{R(e)S[T(e)R(p)]\})]\}) \\ &= T\{R(e)S[T(p)R(T(e)S\{R(e)S[T(e)R(p)]\})]\} \\ &= T[R(e)S(T(p)R\{T(e)S[T(e)R(p)]\})] \\ &= T(R(e)S\{T(p)R[T(e)R(p)]\}) = T\{R(e)S[T(p)R(p)]\} \\ &= T[R(e)S(p)], \end{aligned}$$

it follows from (5.2) that

$$x(p)y(p) = G(R/S; e)r(R/S; T(e) - T[R(e)S(p)]);$$

and the points p and q lie on the same S -line if, and only if, $x(p)y(p) = x(q)y(q)$.

Since two points of the net N are equal if, and only if, they lie on the same R -line, S -line, and T -line, it follows that the correspondence which maps p upon $(x(p), y(p))$ is an isomorphism between the two nets.

Since finally

$$T\{R(e)S[T(e)R(e)]\} = T[R(e)S(e)] = T(e),$$

it follows that $(x(e), y(e)) = (1, 1)$, and this completes the proof.

The following remark may be added to the proof. Under the anti-isomorphism, considered in Theorem 5.3, the coordinate $x(p)$ is mapped upon

$$G(T/S; e)r(T/S; R(e) - R(p)),$$

as follows from Theorem 5.3 (a), and since the transformations in this coset map e upon $T(e)R(p)$, it follows that the transformations in

$$G(T/S; e)r(T/S; R(e) - R(p))G(R/S; e)r(R/S; T(e) - T(p)),$$

defined in the customary sense of group theory, map e upon p .

7. The uniqueness theorem. In the last section it has been shown that every net may be represented in the form $N(M)$, where M is a division system with unit, and that every division system M determines a net $N(M)$.

THEOREM 7.1. *The nets $N(M)$ and $N(L)$, derived from the division systems M and L , both of them with unit, are isomorphic if, and only if, M and L are similar systems.*

It ought to be remembered that isomorphisms map R -lines upon R -lines, S -lines upon S -lines, and T -lines upon T -lines.

Proof. Put $E = N(M)$, $F = N(L)$, $e = (1, 1)$ in E , $f = (1, 1)$ in F . To avoid confusion we denote the transformations $r(R/S; X - Y)$ defined for the nets E and F by $r(E; R/S; X - Y)$ and $r(F; R/S; X - Y)$, respectively, and the other notations are amplified accordingly.

It is a consequence of Theorem 6.2 that we may write without loss in generality $M = M(E; R/S; e)$, $L = M(F; R/S; f)$.

If there exists an isomorphism κ of E upon F , then $M(E; R/S; e)$ and $M(F; R/S; e^*)$ are isomorphic. $M(F; R/S; e^*)$ and $M(F; R/S; f)$ are similar by Theorem 5.1 (b). Consequently, M and L are similar.

If conversely M and L are similar, then it follows from Theorem 5.1 (b) and $L = M(F; R/S; f)$ that there exists a point e' in F so that M and $M(F; R/S; e')$ are isomorphic. Then it follows from Theorem 6.3 that there exists an isomorphism of F upon the net $N[M(F; R/S; e')]$ which maps e' upon the point $(1, 1)$ of this latter net. But M and $M(F; R/S; e')$ being isomorphic, it follows now that there exists an isomorphism of F upon E which maps e' upon e , and this completes the proof.

The net E determines uniquely the class $M(E; R/S)$ of all the systems $M(E; R/S; p)$ for p in E , and this class $M(E; R/S)$ is just a complete class of similar division systems with unit. As such it is completely determined by each of its individual members.

A class of similar division systems with unit determines uniquely, and is in its turn determined uniquely by, (a) a group G , (b) a class C of conjugate subgroups of G whose crosscut is 1, (c) a class D of "similar" sets of representatives of the right cosets of G modulo the subgroups in C .

In our case, $G = G(E; R/S)$, C is the class of subgroups $G(E; R/S; p)$, and D consists of the sets of transformations $r(E; R/S; X - Y)$ for fixed X and variable Y , this latter class being termed $D(E; R/S)$.

Now the Theorem 7.1 may be stated, as a corollary, in the following form:

COROLLARY 7.2. *The nets E and F are isomorphic if, and only if, there exists an isomorphism of the group $G(E; R/S)$ upon the group $G(F; R/S)$ which maps $C(E; R/S)$ upon $C(F; R/S)$ and $D(E; R/S)$ upon $D(F; R/S)$.*

Thus it may be stated as a summary of the results in §§6, 7 that the theory of nets is the same as the theory of classes of similar division systems

with unit, and that it is the same as the theory of a group plus a complete set of conjugate subgroups whose crosscut is 1 plus a class of similar sets of representatives of the right cosets of the group modulo these subgroups.

Finally it may be noted that the proof of Theorem 7.1 contains a proof of the following assertion:

COROLLARY 7.3. *There exists an automorphism of the net which maps the point p upon the point q if, and only if, $M(R/S; p)$ and $M(R/S; q)$ are isomorphic.*

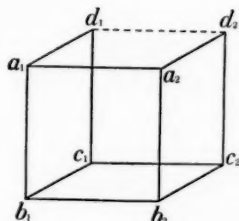


FIG. 5

8. Group-nets. If M is a group, then the net $N(M)$, in the terminology of §6, may be termed a *group-net*. These group-nets have furnished the historical starting point of the theory of nets. To characterize the group-nets, the following net property has been introduced.*

PROPERTY R-S. *If the points a_i, b_i, c_i, d_i form a parallelogram, that is, if $R(a_i) = R(b_i)$, $R(c_i) = R(d_i)$, $S(a_i) = S(d_i)$, $S(b_i) = S(c_i)$, and if $T(a_1) = T(a_2)$, $T(b_1) = T(b_2)$, $T(c_1) = T(c_2)$, that is, if three of the vertices are perspective, then $T(d_1) = T(d_2)$.*

Property R-S is clearly symmetric in R and S , and it is illustrated by Fig. 5.

THEOREM 8.1. *The following properties of a net are equivalent:*

- (1) *The net is a group-net.*
- (2) *An anti-isomorphism of $M(R/S; e)$ upon $M(S/R; e)$ is defined in mapping $G(R/S; e)r(R/S; T(e) - X)$ upon $G(S/R; e)r(S/R; T(e) - X)$.*
- (3) *The net has Property R-S.*
- (4) $G(R/S) = G(R/T)$.
- (5) $M(R/S; e)$ is a group.

* Compare the papers of Thomsen, Kneser, and Reidemeister, mentioned above. We do not state this property in its customary symmetric form, since this weaker asymmetric form is more convenient for our treatment and the stronger symmetric property is a consequence of it; compare Corollary 8.2 below.

$$r(S/R; T(e) - Y)r(S/R; T(e) - X)$$

maps e upon the point $S(e)R(XS\{T(e)R[S(e)Y]\})$. Hence it follows from condition (2) that these two points lie on the same T -line. If now the points b_i of Property R - S are, in particular, points on $T(e)$, a_2 is on $S(e)$ and c_1 on $R(e)$, then in choosing $X = T(c_1)$ and $Y = T(a_1)$, we find

$$d_1 = R(e)S(YR\{T(e)S[R(e)X]\}), \quad d_2 = S(e)R(XS\{T(e)R[S(e)Y]\}),$$

and it follows now from what has been proved that Property R - S holds true at least if the points a_2, b_1, b_2, c_1 are in the special position indicated above.

To derive the general R - S property from the special one, one proves, as indicated in Fig. 7, that the points h_1 and h_2 as well as h_2 and h_3 are on the

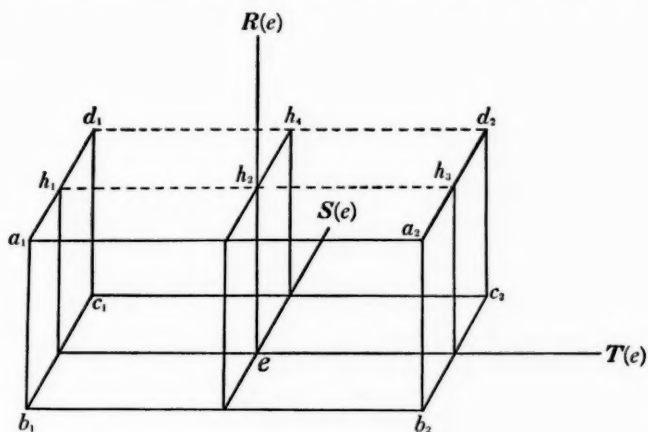


FIG. 7

same T -line as a consequence of the special R - S property, and that therefore both d_1 and h_4 as well as h_4 and d_2 are on the same T -line; this proves quite generally that (3) is a consequence of (2).

Suppose now that the net has Property R - S , that X and Y are two T -lines, and that the points c_1 and c_2 are on the same T -line. Put $b_i = S(c_i)X$, $a_i = R(b_i)Y$, and $d_i = R(c_i)S(a_i)$. Then it is a consequence of the R - S property that d_1 and d_2 lie on the same T -line. But it is a consequence of Theorem 4.3 that $r(R/S; X - Y)$ maps c_i upon d_i . Hence all the transformations $r(R/S; X - Y)$ map T -lines upon T -lines and are therefore at the same time R/T -transformations. This implies, by Theorem 4.3 and Corollary 4.5, that the transformations $r(R/T; X - Y)$ are R/S -transformations, and this proves that (4) is a consequence of (3).

Assume now that (4) is satisfied and that w is some element in $G(R/S; e)$. Then w maps every R -line upon itself, every S -line upon an S -line, and every T -line upon a T -line. In particular, therefore, w maps $T(e)$ upon itself. Hence it follows from Theorem 4.3 that

$$w = r(R/S; T(e) - T(e)) = 1.$$

This shows that $G(R/S; e) = 1$, and this proves that $M(R/S; e)$ is a group; that is, that (5) is a consequence of (4).

That finally (1) is a consequence of (5), is a consequence of Theorem 6.3.

COROLLARY 8.2. *If a net satisfies the conditions (1) to (5) of Theorem 8.1, then*

(i) $G(R/S) = G(R/T)$, $G(S/T) = G(S/R)$, $G(T/R) = G(T/S)$ are isomorphic groups;

(ii) *The net has the R - S , the S - T , and the T - R properties.*

Proof. If $M(R/S; e)$ is a group, then it follows from Theorem 5.3 that $M(T/S; e)$ is an isomorphic group. Hence the net has the S - T property and therefore has the T - R property too. Since $M(U/V; e)$ is a group, it follows that $M(U/V; e) = G(U/V)$, and from the above statement it follows that all these groups are isomorphic. The equalities in (i) are now consequences of (ii) and Theorem 8.1.

COROLLARY 8.3. *Group-nets are isomorphic if, and only if, they are derived from isomorphic groups.*

This is a consequence of Theorem 8.1, Corollary 8.2, and Theorem 7.1.

If one is only interested in the proof of Thomsen's theorem, that is, in the equivalence of the assertions (1), (3), (4), and (5) of Theorem 8.1, then the proof can be simplified very much, since a simple calculation shows that (3) is a consequence of (1).*

One might miss here the symmetry of Kneser's treatment of this theory. But the assertion (2) of Theorem 8.1 makes it probable that such a symmetric treatment will only be possible in restricted cases.† As a matter of fact, it seems to be an interesting problem to investigate symmetry properties of the nets and their relation to the group-theoretical representation of the nets.

An R/S -transformation of the net is an automorphism of the net if, and only if, it is at the same time an R/T -transformation. The R/S -transformations which are net automorphisms are certainly all of the form $r(R/S; X - Y)$, and thus it may be said that *the crosscut of $G(R/S)$ and $G(R/T)$ consists of exactly those R/S -transformations which are net automorphisms.*

* Cf., for example, Kneser, op. cit., p. 148.

† Cf. Bol, op. cit., §3, where a symmetric treatment of the quasi-group-nets is given.

It is now easy to verify that the conditions (1) to (5) of Theorem 8.1 are equivalent to the following condition:

(6) *The crosscut of $G(R/S)$ and $G(R/T)$ is a transitive group of permutations of the T -lines.*

9. Nets and simply transitive systems of permutations. It has been pointed out in §3 that the theory of classes of similar division systems with unit is completely equivalent to the theory of classes of similar simply transitive systems of permutations containing the identity; and it has been proved in §§6, 7 that the theory of nets is equivalent to the theory of classes of similar division systems with unit. The theory of nets is therefore equivalent to the theory of classes of similar simply transitive systems of permutations containing the identity. To put the concrete significance of this abstract equivalence into evidence is the object of this section.

If N is a net, and if E is a T -line in N , then the transformations $r(R/S; E-X)$ form a system $P(R/S; E) = P(N; R/S; E)$ of permutations of the S -lines of the net.

(9.1) (a) *$P(R/S; E)$ is a simply transitive system of permutations which contains the identity.*

(b) *If E and F are two T -lines, then $P(R/S; E)$ and $P(R/S; F)$ are similar.*

Proof. The first of these facts is a consequence of Corollary 4.5 and of $r(R/S; E-E) = 1$. The second of these facts is a consequence of Theorem 3.5 and of Theorem 5.1, since $P(R/S; E)$ is isomorphic to the system $P[M(R/S; e)]$ of the right translations of the division system $M(R/S; e)$, provided e is a point on E .

If P is a simply transitive system of permutations of the elements in the set Q , and if P contains the identity, then a net $N'(P)$ may be derived from P in the following fashion. The points of this net are the pairs (q, p) for q in Q and p in P . There corresponds furthermore an R -line as well as an S -line to every element in Q , whereas to every element in P there corresponds a T -line. The point (q, p) lies finally (a) on the R -line corresponding to q , (b) on the S -line corresponding to q^p , (c) on the T -line corresponding to p . $N'(P)$ is a net since P is simply transitive. That P contains the identity is not needed for this inference.

(9.2) *If P is a simply transitive system of permutations which contains the identity and if E is the T -line in the net $N'(P)$ which corresponds to the identity in P , then P and $P[N'(P); R/S; E]$ are isomorphic systems.*

Proof. Suppose that X is the T -line in our net which corresponds to the element x of P . Then it is a consequence of Theorem 4.3 that $r(E-X)$ maps

the point $n = (q, p)$ upon the point $R(n)S\{XR[ES(n)]\}$. Since $S(n)$ is the S -line corresponding to q^p , it follows that $ES(n) = (q^p, 1)$. Consequently we have

$$XR[ES(n)] = (q^p, x), \quad R(n)S\{XR[ES(n)]\} = (q, \widehat{px})$$

where \widehat{px} is the uniquely determined permutation in P which maps the element q onto the element $(q^p)^x$. This shows that $r(E-X)$ maps the S -line corresponding to q^p upon the S -line corresponding to $(q^p)^x$, so that the permutation x of the elements in the set Q and the permutation $r(E-X)$ of the S -lines of our net are essentially the same, and this proves our statement.

THEOREM 9.3. *The nets N and $N'[P(N; R/S; E)]$ are isomorphic (for every T -line E of the net N).*

Proof. If n is any point of the net N , then put

$$q(n) = S[ER(n)], \quad p(n) = r(R/S; E - T(n)).$$

This is a single-valued transformation mapping the point n of the net N upon the point $(q(n), p(n))$ of the net $N'[P(N; R/S; E)] = N'$. The points n and m are on the same R -line if, and only if, $q(n) = q(m)$; and they are on the same T -line if, and only if, $p(n) = p(m)$. This implies, in particular, that the correspondence is a one-one correspondence. If (q, p) is some point of N' , then q is an S -line and $p = r(R/S; E - X)$ for some T -line X . The point $n = XR(Eq)$ satisfies clearly $p(n) = p$, and it satisfies $q(n) = q$, since $ER(Eq) = Eq$ and $q(n) = S[ER(Eq)] = S(Eq) = q$. Our correspondence maps, therefore, the net N upon the whole net N' . The transformation $p(n)$ maps the point $ER(n)$ upon the point $T(n)R(n) = n$ and therefore the S -line $q(n)$ upon the S -line $S(n)$. The points n and m are therefore on the same S -line if, and only if, $q(n)^{p(n)} = q(m)^{p(m)}$, and this completes the proof that the nets N and N' are isomorphic.

THEOREM 9.4. *Suppose that E is a T -line of the net N , and that E^* is a T -line of the net N^* . The nets N and N^* are isomorphic if, and only if, $P(N; R/S; E)$ and $P(N^*; R/S; E^*)$ are similar systems.*

Proof. Denote by e some point on the T -line E and by f some point on the T -line E^* . Then it is a consequence of Corollary 7.2 that the nets N and N^* are isomorphic if, and only if, the division systems $M(N; R/S; e)$ and $M(N^*; R/S; f)$ are similar; and it is a consequence of Theorem 3.5 that these two division systems are similar if, and only if, $P[M(N; R/S; e)]$ and $P[M(N^*; R/S; f)]$ are similar systems of permutations. But this proves our statement, since the first of these systems of permutations is isomorphic with $P(N; R/S; E)$ and the second one with $P(N^*; R/S; E^*)$.

REMARK 9.5. *It is very simple indeed to derive Theorem 9.3 from Theorem 9.4. To do this one has only to remark that as a consequence of (9.2) the systems $P(N; R/S; E)$ and $P\{N'[P(N; R/S; E)]; R/S; E'\}$ are isomorphic systems of permutations.*

COROLLARY 9.6. *$N'(P)$ and $N'(P^*)$ are isomorphic if, and only if, the systems P and P^* of permutations are similar.*

This is a consequence of (9.2) and Theorem 9.4.

Finally it should be pointed out that the treatment given in this section is somewhat more symmetric than the one outlined in §§6, 7. For here we had to give preference to some T -line E , whereas in the former case we had to distinguish a certain point e .

10. **Subnets.** A system K of points, R -lines, S -lines, and T -lines of a net N is termed a *subnet* of N , if K is a net under the incidence relations, as defined in N .

The crosscut of a system of subnets of a net is either empty or itself a subnet. There exists, consequently, corresponding to every configuration in a net a smallest containing subnet.

LEMMA 10.1. *If K is a subnet of the net N , if E and X are two T -lines in K , then $r(R/S; E-X)$ maps K upon itself.*

Proof. $r(R/S; E-X)$ maps, by Theorem 4.3, the point p of K upon the point $R(p)S\{XR[ES(p)]\}$. The line $S(p)$ is in K , as p is in K ; and $ES(p)$ belongs to K , since E is in K . This implies that $R[ES(p)]$ belongs to K ; and as X is in K , both $XR[ES(p)]$ and $S\{XR[ES(p)]\}$ belong to K . The line $R(p)$ belongs to K , since p is a point in K ; and it thus follows finally that $r(R/S; E-X)$ maps every point of K upon a point of K . Since

$$r(R/S; E-X)^{-1} = r(R/S; X-E),$$

it follows that the inverse of $r(R/S; E-X)$ maps every point of K upon a point of K ; and this proves that $r(R/S; E-X)$ maps K upon itself.

COROLLARY 10.2. *If K is a subnet of the net N , E a fixed T -line of K , and X a variable T -line in K , then the transformations $r(R/S; E-X)$ form a simply transitive system of permutations of the S -lines in K which contains the identity.*

This is a consequence of Lemma 10.1 and Corollary 4.5.

THEOREM 10.3. *Two subnets are identical if they have all S -lines and one T -line in common.*

REMARK. *Note that two subnets have one point in common if, and only if, they have one S -line and one T -line in common.*

Proof. Suppose that the two subnets have the T -line E in common. The set of all the transformations $r(R/S; E-X)$ is, by Corollary 4.5, simply transitive on the set of all the S -lines of the whole net. It contains, therefore, at most one subset which is simply transitive on a given set of S -lines. Thus it follows from Corollary 10.2 that the two subnets have all the T -lines in common. But they then consist of the same points too and are therefore equal.

THEOREM 10.4. *If Z is a set of S -lines and D a set of T -lines, and if E is a T -line in D so that the transformations $r(R/S; E-X)$ for X in D form a simply transitive system of permutations of the S -lines in Z , then there exists one (and only one) subset K whose set of S -lines is Z and whose set of T -lines is D . The points of K are exactly the points UV for U in Z and V in D , and the lines $R(UV)$ are its R -lines.*

Proof. During the proof of Theorem 9.3 it has been shown that the net may be represented in the form $N'(P)$, if one only represents the point n of the net by the coordinates $q(n) = S[ER(n)]$, $p(n) = r(R/S; E-T(n))$. Denote now by K the set of all those points (q, p) whose coordinates satisfy the condition that q is in Z and p is in D . It is a consequence of Theorem 9.3 that these points form a net $N'(P^*)$, since the set P^* of the permutations $r(R/S; E-X)$, for X in D , contains the identity and is simply transitive on the set Z of S -lines. Thus K is a subnet of our net. The S -line through the point (q, p) is just the line q^p which belongs to Z , since q belongs to Z and since the p in P^* map Z upon itself. If W is an S -line in Z , p any element in P^* , then there exists one and only one q in Z , so that $q^p = W$ and $W = S[(q, p)]$ belongs therefore to K . Thus the S -lines of K form exactly the set Z . That D is just the set of the T -lines in K , is obvious, and this completes the proof.

We add another characterization of the subnets of a net. Here we make use of the fact that every net may be represented in the form $N(M)$, where M is a division system with unit and where the point $(1, 1)$ may be prescribed at random, choosing $M = M(R/S; e)$. For this characterization we shall need the following concept: If M is a division system with unit, then the subset Q of M is said to be *closed* if Q is a division system with unit under the multiplication, as defined in M . Consequently, the subset Q of M is closed if, and only if, Q contains the unit and contains with the elements u and v also uv and the elements x and y , satisfying $ux = v$ and $yu = v$.

THEOREM 10.5. *The set U of points in the net $N(M)$, where M is a division system with unit, forms together with the R -lines, S -lines, and T -lines through points of U a subnet of $N(M)$ which contains the point $(1, 1)$ if, and only if, there exists a closed subset Q of M so that U consists exactly of the points (a, b) for a and b in Q .*

Proof. The sufficiency of the condition is a consequence of Theorem 6.1. Assume now that the set W , consisting of the points in U and of the R -lines, S -lines, and T -lines through points in U , is a subnet of $N(M)$ which contains the point $(1, 1)$. Then denote by Q the set of all those elements u in M so that $(u, 1)$ is in U .

(10.5.1) *If (u, v) belongs to U , then u and v belong to Q .*

There belong to W certainly the R -line, the S -line, and the T -line, corresponding to 1, since $(1, 1)$ is in U . Since (u, v) is in U , there belongs to W the R -line corresponding to u , the S -line corresponding to uv , and the T -line corresponding to v . Hence $(u, 1)$ and $(1, v)$ are in U and u is in Q . As $(1, v)$ is in U , the S -line corresponding to v is in W , and $(v, 1)$ is therefore in U , v in Q .

(10.5.2) *If u and v belong to Q , then (u, v) belongs to U .*

If u and v belong to Q , then $(u, 1)$ and $(v, 1)$ belong to U . Hence the R -line corresponding to u and the S -line corresponding to v are in W . As W contains the R -line corresponding to 1, it follows that U contains $(1, v)$, and therefore that W contains the T -line corresponding to v . Since W contains the R -line corresponding to u , and the T -line corresponding to v , it follows that U contains (u, v) .

(10.5.3) *Q is closed in M .*

The unit 1 is in Q , since $(1, 1)$ is in U . If u and v are in Q , then (u, v) is in U by (10.5.2). W contains therefore the S -line corresponding to uv . Since W contains the T -line corresponding to 1, $(uv, 1)$ is in U and uv is in Q . Since v is in Q , $(v, 1)$ is in U , and the S -line corresponding to v is in W . Since u is in Q , $(u, 1)$ is in U , and the R -line corresponding to u is in W . Hence there is in U a point (u, x) so that $ux = v$; and it follows from (10.5.1) that the solution x of $ux = v$ is in Q . Since 1 and u are in Q , it follows from (10.5.2) that $(1, u)$ is in U and the T -line corresponding to u is in W . Since the S -line corresponding to v is in W , as remarked before, there is in U a point (y, u) for $yu = v$ and it follows from (10.5.1) that the solution y of $yu = v$ is in Q . This proves (10.5.3).

It is a consequence of (10.5.1) and (10.5.2) that U is exactly the set of the points (a, b) for a and b in Q ; and as Q is, by (10.5.3), closed in M , this completes the proof of the theorem.

If we use the term "lattice," the above result may be stated in the following form: If e is a point of the net N , $L(e)$ the lattice of all the subnets of N which contain e , then $L(e)$ and the lattice of the closed subsets in $M(R/S; e)$ are isomorphic.

UNIVERSITY OF ILLINOIS,
URBANA, ILL.

ON A GENERALIZATION OF THE STIELTJES MOMENT PROBLEM*

BY

R. P. BOAS, JR.†

Introduction. The moment problem of Stieltjes is the problem of determining the non-decreasing solutions $\alpha(t)$ of the set of equations

$$(1) \quad \mu_n = \int_0^\infty t^n d\alpha(t), \quad n = 0, 1, 2, \dots;$$

the phrase "a moment problem" is also used to describe the system (1) itself. If a solution of (1) is known, there arises the further question of whether or not the function $\alpha(t)$ is unique.‡ It is this question which we shall discuss for a generalized moment problem, namely

$$(2) \quad \mu_n = \int_0^\infty t^{\lambda_n} d\alpha(t), \quad 0 = \lambda_0 < \lambda_1 < \dots < \lambda_n \rightarrow \infty.$$

If (2) has a unique solution $\alpha(t)$, we say that (2) is determined; otherwise (2) is said to be undetermined.

The various classical methods for the study of (1) seem not to apply to (2), since they depend too much on special properties of the sequence $\{\lambda_n\} = \{n\}$. We shall discuss the determination problem for (2) by considering the function

$$(3) \quad f(z) = \int_0^\infty t^z d\alpha(t),$$

which is analytic for $\Re(z) > 0$, and takes the values μ_n at the points λ_n ; since $\alpha(t)$ is non-decreasing, the growth of $f(z)$ is governed by the growth of the μ_n . We obtain sufficient conditions for (2) to be determined by applying a fundamental theorem of T. Carleman concerning the growth of functions analytic in a half-plane.

The criteria obtained in this way are probably not the best possible; when $\lambda_n = n$, they are certainly not, since we obtain

* Presented to the Society, April 15, 1938; received by the editors January 21, 1939.

† National Research Fellow.

‡ Two solutions $\alpha(t)$ of (1) are considered the same if they have the same "normalization," determined by $\alpha(0) = 0$, $\alpha(t) = [\alpha(t+) + \alpha(t-)]/2$, $t > 0$.

$$(4) \quad \mu_n^{1/(2n)} = o(n)$$

as a sufficient condition for (1) to be determined; this is much weaker than Carleman's criterion, $\sum_{n=1}^{\infty} \mu_n^{-1/(2n)}$ divergent. On the other hand, we obtain what may be regarded as new criteria for the case $\lambda_n = n$, since we shall show that (4) is still a sufficient condition for (1) to be determined if we disregard a set of integers n_k such that

$$\sum_{k=1}^{\infty} \frac{1}{n_k} < \infty, \quad \liminf_{r \rightarrow \infty} \frac{\Delta(r)}{r^{1/2}} < \infty,$$

where $\Delta(r)$ is the maximum number of consecutive integers which are n_k 's for $n_k \leq r$ (for example, $n_k = k^2$).

Another interesting case is that where

$$\limsup_{n \rightarrow \infty} |\lambda_n - n| < \infty.$$

In this case, (4) is again a sufficient condition for (2) to be determined.

In general, the denser the λ_n , the less we have to restrict the growth of the μ_n to be sure that (2) will be determined. On the other hand, if the λ_n are so sparse that $\sum_{n=1}^{\infty} 1/\lambda_n < \infty$, there are presumably no criteria for determination depending only on the order of magnitude of the μ_n . For, since even the moment problem for a finite interval,

$$(5) \quad \mu_n = \int_0^1 t^{\lambda_n} d\alpha(t),$$

may be undetermined in this case,* we could only hope to show that (2) would be determined if the μ_n approached zero with extreme rapidity. But, if $\alpha(t)$ has a point of increase $t_0 > 0$, we necessarily have

$$\mu_n \geq t_0^{\lambda_n} [\alpha(\infty) - \alpha(t_0 -)],$$

and so a lower limit to the rate of decrease of the μ_n .

1. Let

$$(1.1) \quad \lambda_0 = 0, \quad \lambda_1 \geq 1, \quad \lambda_n \uparrow \infty \text{ as } n \uparrow \infty.$$

We write

* In fact, if $d\alpha(t) = a(t)dt$ and $a(t) \geq \delta > 0$, ($0 \leq t \leq 1$), then (5) is undetermined. For, by Müntz's theorem (see, for example, R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. 19, 1934, p. 36), there is a continuous $b(t)$ such that $\int_0^1 t^{\lambda_n} b(t) dt = 0$, ($n=0, 1, 2, \dots$), and $b(t) \not\equiv 0$. We may suppose $b(t) \leq \delta$; then $\int_0^1 t^{\lambda_n} [a(t) - b(t)] dt = \int_0^1 t^{\lambda_n} a(t) dt$, and $a(t) - b(t) \geq 0$. Cf. F. Hallenbach, *Zur Theorie der Limitierungsverfahren von Doppelfolgen*, Dissertation, Bonn, 1933, p. 94.

$$(1.2) \quad d(\lambda_n) = \lambda_n - \lambda_{n-1}; \quad \Delta(r) = \max_{\lambda_n \leq r} d(\lambda_n);$$

$$(1.3) \quad \xi_0 = 0, \quad \xi_n = \lambda_n - 1, \quad n \geq 1;$$

and $z = x + iy = re^{i\theta}$. Throughout the paper, A denotes a constant, depending on the data of the problem in hand, and not necessarily the same at each appearance.

In this section we estimate the expression

$$(1.4) \quad M(r) = \frac{1}{2} \int_{-\pi/2}^{\pi/2} \log |f(re^{i\theta})| \cos \theta \, d\theta,$$

formed with a function $f(z)$, analytic for $x \geq 0$, and subject to a limitation of the form

$$(1.5) \quad |f(x + iy)| \leq A\mu_n, \quad \xi_{n-1} < x \leq \xi_n, \quad n = 1, 2, \dots;$$

we suppose that $\mu_n \geq A > 0$, or (without loss of generality) $\mu_n \geq 1$. In the applications to moment problems, the μ_n and λ_n will be the μ_n and λ_n of the introduction, and $f(z)$ will be essentially the difference of the functions (3) formed for two solutions of the moment problem under consideration. The relevance of the expression (1.4) is clear from inspection of Carleman's theorem (quoted in §2).

THEOREM 1. *Let $f(z)$ be analytic for $x \geq 0$, let $f(z)$ satisfy (1.4), and let*

$$(1.6) \quad 0 \leq \log \mu_n \leq 2\lambda_n G(\lambda_n),$$

where $G(r)$ is a non-decreasing function. Then for any $\epsilon > 0$ there is an m_ϵ such that for $m > m_\epsilon$

$$(1.7) \quad \frac{M(\xi_m)}{\xi_m} \leq \frac{A}{\lambda_m} + G(\lambda_m) \left\{ \frac{\pi}{2} + 2^{3/2}(1 + \epsilon)d(\lambda_m)^{1/2}\lambda_m^{-1/2} + 2^{1/2}(1 + \epsilon)\Delta(\lambda_m)d(\lambda_m)^{-1/2}\lambda_m^{-1/2} \right\}.$$

We have, from (1.5),

$$(1.8) \quad \log |f(re^{i\theta})| \leq A + \log \mu_n, \quad \xi_{n-1} < r \cos \theta \leq \xi_n.$$

We then have

$$(1.9) \quad \begin{aligned} \frac{M(\xi_m)}{\xi_m} &= \frac{1}{\xi_m} \int_0^{\phi_{m-1}} (A + \log \mu_m) \cos \theta \, d\theta \\ &\quad + \frac{1}{\xi_m} \sum_{k=1}^{m-1} \int_{\phi_k}^{\phi_{k-1}} (A + \log \mu_k) \cos \theta \, d\theta, \end{aligned}$$

where $\phi_k = \cos^{-1}(\xi_k/\xi_m)$, ($k=0, 1, \dots, m-1$). That is,

$$\begin{aligned}\frac{M(\xi_m)}{\xi_m} &\leq \frac{A}{\xi_m} + \frac{\log \mu_m}{\xi_m} \left(1 - \frac{\xi_{m-1}^2}{\xi_m^2}\right)^{1/2} \\ &\quad + \frac{1}{\xi_m} \sum_{k=1}^{m-1} \left\{ \left(1 - \frac{\xi_{k-1}^2}{\xi_m^2}\right)^{1/2} - \left(1 - \frac{\xi_k^2}{\xi_m^2}\right)^{1/2} \right\} \log \mu_k \\ &= \frac{A}{\xi_m} + P_m + 2 \sum_{k=1}^{m-1} R_k \log \mu_k,\end{aligned}$$

say.

Using (1.6), we have

$$\begin{aligned}P_m &\leq 2\lambda_m \xi_m^{-2} G(\lambda_m) (\xi_m^2 - \xi_{m-1}^2)^{1/2} \leq 2^{3/2} \lambda_m \xi_m^{-3/2} G(\lambda_m) (\xi_m - \xi_{m-1})^{1/2} \\ &= 2^{3/2} \lambda_m (\lambda_m - 1)^{-3/2} G(\lambda_m) d(\lambda_m)^{1/2} \leq 2^{3/2} (1 + \epsilon) \lambda_m^{-1/2} G(\lambda_m) d(\lambda_m)^{1/2},\end{aligned}$$

for m sufficiently large.

Again,

$$R_k \log \mu_k \leq \frac{1}{2} \frac{\xi_k d(\lambda_k) \log \mu_k}{\xi_m^2 (\xi_m^2 - \xi_k^2)^{1/2}} \leq \frac{\lambda_k \xi_k G(\lambda_k) d(\lambda_k)}{\xi_m^2 (\xi_m^2 - \xi_k^2)^{1/2}},$$

and since $G(r)$ is non-decreasing,

$$\sum_{k=1}^{m-1} R_k \log \mu_k \leq \frac{G(\lambda_m)}{\xi_m^2} \sum_{k=1}^{m-1} g(\xi_k) (\xi_k - \xi_{k-1}),$$

where $g(x) = x(x-1)(\xi_m^2 - x^2)^{-1/2}$.

Now we have*

$$\begin{aligned}\sum_{k=1}^{m-1} g(\xi_k) (\xi_k - \xi_{k-1}) &= \xi_{m-1} g(\xi_{m-1}) - \sum_{k=0}^{m-2} \xi_k [g(\xi_{k+1}) - g(\xi_k)] \\ &= \xi_{m-1} g(\xi_{m-1}) - \sum_{k=0}^{m-2} \int_{\xi_k}^{\xi_{k+1}} \xi_k g'(x) dx \\ &= \xi_{m-1} g(\xi_{m-1}) - \int_0^{\xi_{m-1}} \langle x \rangle g'(x) dx,\end{aligned}$$

where $\langle x \rangle$ denotes the largest ξ_k not exceeding x . Since

$$\int_0^{\xi_{m-1}} x g'(x) dx = \xi_{m-1} g(\xi_{m-1}) - \int_0^{\xi_{m-1}} g(x) dx,$$

we obtain

* Cf. the derivation of Euler's summation formula: K. Knopp, *Theory and Application of Infinite Series*, 1928, p. 522.

$$\sum_{k=1}^{m-1} g(\xi_k)(\xi_k - \xi_{k-1}) = \int_0^{\xi_{m-1}} g(x)dx + \int_0^{\xi_m} (x - \langle x \rangle)g'(x)dx.$$

Now

$$\int_0^{\xi_{m-1}} g(x)dx = \int_0^{\xi_{m-1}} \frac{x(x-1)dx}{(\xi_m^2 - x^2)^{1/2}} \leq \int_0^{\xi_m} \frac{x^2 dx}{(\xi_m^2 - x^2)^{1/2}} = \frac{\pi \xi_m^2}{4}.$$

And, since $g(x)$ is an increasing function,

$$\left| \int_0^{\xi_{m-1}} (x - \langle x \rangle)g'(x)dx \right| \leq \frac{\xi_{m-1}\lambda_{m-1}\Delta(\xi_m)}{(\xi_m^2 - \xi_{m-1}^2)^{1/2}} \leq \frac{(1+\epsilon)\xi_m^{3/2}\Delta(\xi_m)}{2^{1/2}d(\lambda_m)^{1/2}},$$

for m sufficiently large.

Collecting results, we obtain

$$\frac{M(\xi_m)}{\xi_m} \leq \frac{A}{\xi_m} + \frac{2^{3/2}(1+\epsilon)G(\lambda_m)d(\lambda_m)^{1/2}}{\lambda_m^{1/2}} + G(\lambda_m) \left\{ \frac{\pi}{2} + \frac{2^{1/2}(1+\epsilon)\Delta(\xi_m)}{d(\lambda_m)^{1/2}\xi_m^{1/2}} \right\},$$

for sufficiently large m ; this is (1.7).

2. We now consider the moment problem

$$(2.1) \quad \mu_n = \int_0^\infty t^{n\alpha} d\alpha(t),$$

where $\alpha(t)$ is non-decreasing, $\lambda_0=0$, $\lambda_1 \geq 1$, $\lambda_n \uparrow \infty$, and

$$(2.2) \quad \sum_{n=1}^\infty \frac{1}{\lambda_n} \text{ diverges.}$$

We may then suppose that $\mu_n \rightarrow \infty$, since otherwise $\alpha(t)$ would be constant outside $(0, 1)$, and (2.1) would be determined.* Hence we may (and shall) suppose that $\mu_n \geq 1$, ($n=0, 1, 2, \dots$).

It is reasonable to suppose that the μ_n satisfy an inequality of the form

$$(2.3) \quad \mu_n^{1/(2\lambda_n)} \leq e^{G(\lambda_n)}, \quad G(r) \uparrow \infty \text{ as } r \uparrow \infty;$$

or, more conveniently written,

$$(2.4) \quad \log \mu_n \leq 2\lambda_n G(\lambda_n).$$

We define the expression $Q(r)$ by

$$(2.5) \quad Q(r) = \sum_{\lambda_n \leq r} \left(\frac{1}{\lambda_n} - \frac{\lambda_n}{r^2} \right),$$

* F. Hausdorff, *Summationsmethoden und Momentfolgen*, II, *Mathematische Zeitschrift*, vol. 9 (1921), pp. 280-299; p. 287.

and define $d(\lambda_n)$, $\Delta(r)$, and ξ_n by relations (1.2), (1.3). We shall prove

THEOREM 2. *If (2.1) is undetermined, then for any $\epsilon > 0$ and m sufficiently large,*

$$(2.6) \quad Q(\xi_m) \leq A + G(\lambda_m) \left\{ 1 + \frac{C_1 d(\lambda_m)^{1/2}}{\lambda_m^{1/2}} + \frac{C_2 \Delta(\lambda_m)}{\lambda_m^{1/2} d(\lambda_m)^{1/2}} \right\},$$

where* $C_1 = 2^{5/2} \pi^{-1} (1 + \epsilon)$, $C_2 = 2^{3/2} \pi^{-1} (1 + \epsilon)$.

We may state less forbidding special cases of (2.6) if we suppose that the growth of the sequence $\{\lambda_n\}$ is very regular. Thus we have

COROLLARY 2.1. *If $d(\lambda_n)$ increases and $d(\lambda_n) = o(\lambda_n)$, then, if (2.1) is undetermined,*

$$(2.7) \quad Q(\xi_m) \leq G(\lambda_m)(1 + o(1)).$$

COROLLARY 2.2. *If $d(\lambda_n)$ decreases and $d(\lambda_n) = o(1/\lambda_n)$, then, if (2.1) is undetermined, (2.7) holds.*

From Theorem 2 it follows that any condition which makes $G(\lambda_n)$ so small that (2.6) is impossible is a sufficient condition for (2.1) to be determined; in §3 we shall give examples of such conditions for special sequences $\{\lambda_n\}$.

We derive Theorem 2 from the estimate of Theorem 1 applied to

CARLEMAN'S THEOREM.† *Let $f(z)$ be analytic for $x \geq 0$, and let $r_n e^{i\theta_n}$, ($r_1 \leq r_2 \leq \dots$), denote the zeros of $f(z)$ for $x \geq 0$, each counted according to its multiplicity. Then if $R > \rho > 0$,*

$$(2.8) \quad \sum_{\rho < r_n \leq R} \left[\frac{1}{r_n} - \frac{r_n}{R^2} \right] \cos \theta_n = \frac{2M(R)}{\pi R} + A(R) + O(1),$$

where

$$(2.9) \quad A(R) = \frac{1}{2\pi} \int_{\rho}^R \left\{ \frac{1}{r^2} - \frac{1}{R^2} \right\} \log \{ |f(iy)f(-iy)| \} dy,$$

$$M(r) = \frac{1}{2} \int_{-\pi/2}^{\pi/2} \log |f(re^{i\theta})| \cos \theta d\theta,$$

and the term $O(1)$ depends on ρ and is bounded as $R \rightarrow \infty$ for fixed ρ .

Under the hypotheses of Theorem 2, there are two solutions of (2.1); let $\gamma(t)$ be their difference. Consider the function

* The precise values of C_1 and C_2 do not seem very important.

† See, for example, E. C. Titchmarsh, *The Theory of Functions*, 1932, p. 130.

$$(2.10) \quad f(z) = \frac{1}{2} \int_0^\infty t^{z+1} d\gamma(t).$$

Then $f(z)$ is analytic for $x \geq 0$, and has zeros at least at the points $\xi_n = \lambda_n - 1$, ($n=1, 2, \dots$).

Since $\mu_n \rightarrow \infty$, we have for $\xi_{n-1} < x \leq \xi_n$, ($n=0, 1, 2, \dots$),

$$(2.11) \quad |f(x+iy)| \leq \frac{1}{2} \int_0^1 t^{\lambda_{n-1}} |d\gamma(t)| + \frac{1}{2} \int_1^\infty t^{\lambda_n} |d\gamma(t)| \\ \leq \mu_n + A \leq A\mu_n.$$

Now

$$f(iy) = \frac{1}{2} \int_0^\infty t^{1+iy} d\gamma(t), \quad |f(iy)| \leq \frac{1}{2} \int_0^\infty t |d\gamma(t)| \leq \mu_1;$$

consequently $A(R) = O(1)$, $R \rightarrow \infty$.

If we apply Carleman's theorem to $f(z)$, taking $R = \xi_n$ and ρ sufficiently large, use the estimate of Theorem 1 for $M(R)$, and neglect possible zeros of $f(z)$ other than those at the ξ_n (which would only increase the left-hand side of (2.8)), we obtain Theorem 2.

3. We now illustrate Theorem 2 by applying it to a number of specific sequences $\{\lambda_n\}$.

EXAMPLE 1. Let $\lambda_n = n$.

Here $Q(r) = \log r + O(1)$, $r \rightarrow \infty$; $d(\lambda_n) = \Delta(r) \equiv 1$; and (2.6) becomes

$$\log(m-1) \leq A + G(m)(1 + O(m^{-1/2})),$$

which is impossible if $G(r) = \log r + \log \sigma(r)$, where $\liminf_{r \rightarrow \infty} \sigma(r) = 0$. Consequently, the moment problem (2.1) is determined if $\lambda_n = n$ and

$$(3.1) \quad \lim_{n \rightarrow \infty} n^{-1} \frac{1}{\mu_n^{1/(2n)}} = 0.$$

EXAMPLE 2. Let λ_n run through the positive integers with the exception of a set $\{n_k\}$ for which $\sum_{k=1}^\infty 1/n_k < \infty$, and such that $\lim_{r \rightarrow \infty} \Delta(r)r^{-1/2} < \infty$.

Then $Q(r) = \log r + O(1)$, $d(\lambda_n) \geq 1$, and from (2.6) we see that (2.1) is determined if

$$(3.2) \quad \lim_{n \rightarrow \infty} \lambda_n^{-1} \frac{1}{\mu_n^{1/(2\lambda_n)}} = 0.$$

Moreover, as we stated in the introduction, (2.1) is determined even if (3.1) is satisfied. In fact, we may write (3.2) in the form

$$(3.3) \quad \lim_{n \rightarrow \infty} \left(\frac{1}{2n} \cdot \frac{n}{\lambda_n} \log \mu_n - \log n - \log \frac{\lambda_n}{n} \right) = -\infty.$$

If (3.1), or $\lim_{n \rightarrow \infty} [(2n)^{-1} \log \mu_n - \log n] = -\infty$, is satisfied, (3.3) is certainly satisfied if $\lambda_n/n = O(1)$ as $n \rightarrow \infty$. The difference $\lambda_n - n$ is $N(\lambda_n)$, the number of $n_k \leq \lambda_n$; consequently $0 < \delta \leq n/\lambda_n \leq 1$ unless $N(\lambda_n) \sim \lambda_n$, $n \rightarrow \infty$. But if $N(\lambda_n) \geq c\lambda_n$, we have

$$\sum_{n_k \leq \lambda_n} \frac{1}{n_k} \geq \sum_{(1-c)\lambda_n}^{\lambda_n} \frac{1}{k} \sim \log \frac{1}{1-c}, \quad n \rightarrow \infty,$$

so that, since $\sum_{k=1}^{\infty} 1/n_k < \infty$, we must have $N(\lambda_n) \leq c\lambda_n$, $c < 1$, for all sufficiently large n , and hence $\lambda_n/n = O(1)$.

EXAMPLE 3. Let

$$(3.4) \quad |\lambda_n - n| < A, \quad n = 1, 2, \dots$$

Then $Q(r) = \log r + O(1)$, $\limsup_{n \rightarrow \infty} d(\lambda_n) > 0$, and $\Delta(r) \leq 2A$. Consequently, (2.1) is determined if (3.1) is satisfied. Condition (3.4) can, of course, be considerably weakened.

EXAMPLE 4. Let $\lambda_n = n^a$, ($0 < a < 1$).

Then

$$Q(r) = \frac{1}{1-a} r^{(1-a)/a} + O(1), \quad \Delta(\lambda_n) \leq a\rho^{(a-1)/a},$$

$$a\lambda_n^{(a-1)/a}(1 - o(1)) \leq d(\lambda_n) \leq a\lambda_n^{(1-a)/a}.$$

Consequently, for an undetermined moment problem we must have, with $r = \lambda_n = n^a$,

$$(3.5) \quad \frac{1}{1-a} (r-1)^{(1-a)/a} \leq A + G(r) \left\{ 1 + C_1 a^{1/2} r^{-(a^2-a+1)/(2a)} \right. \\ \left. + C_2 a^{1/2} \rho^{(a-1)/a} r^{-(a^2+2a-2)/(2a)} \right\}.*$$

It is clear that we must expect somewhat different results for different values of a .

(i) Let $1 > a > 2(2^{1/2}-1)$, so that $a^2+4a-4 > 0$. If we suppose that

$$(3.6) \quad G(r) \leq (1-a)^{-1} r^{(1-a)/a} + \log \sigma(r), \quad \sigma(r) = o(1),$$

the right-hand side of (3.5) does not exceed

$$A + (1-a)^{-1} r^{(1-a)/a} + \log \sigma(r) + C_1 (1-a)^{-1} a^{1/2} r^{-(a^2-a+1)/(2a)} \\ + C_2 a^{1/2} (1+a)^{-1} \rho^{(a-1)/a} r^{-(a^2+4a-4)/(2a)}.$$

Since $a^2+a-1 > a^2+4a-4 > 0$, if the moment problem is undetermined and $G(r)$ satisfies (3.6), we must have

* We have absorbed the factor $1/(1-o(1))$ into C_2 , which already contained a factor $(1+\epsilon)$, ($\epsilon > 0$).

$$(3.7) \quad (1-a)^{-1}[(r-1)^{(1-a)/a} - r^{(1-a)/a}] \leq \log \sigma(r) + O(1).$$

This, however, is impossible, since the left-hand side of formula (3.7) is $O(r^{(1-2a)/a}) = O(1)$ if $a \geq \frac{1}{2}$. Hence (2.1) is determined if

$$\mu_n^{1/(2n^a)} = o\left\{\exp\left(\frac{n^{1-a}}{1-a}\right)\right\}.$$

(ii) Let $2(2^{1/2}-1) > a > 3^{1/2}-1$, so that $a^2+2a-2 > 0$. If we suppose that

$$G(r) \leq \frac{1-\eta}{1-a} r^{(1-a)/a} + \log \sigma(r), \quad \sigma(r) = o(1),$$

for some $\eta > 0$, the right-hand side of (3.5) does not exceed

$$A + \frac{1-\eta}{1-a} r^{(1-a)/a} + \log \sigma(r) + C_1 \frac{1-\eta}{1-a} a^{1/2} r^{-(a^2+a-1)/(2a)} \\ + C_2 a^{1/2} \frac{1-\eta}{1-a} \rho^{(a-1)/a} r^{-(a^2+4a-4)/(2a)}.$$

Since $(4-4a-a^2)/(2a) < (1-a)/a$ if $a^2+2a-2 > 0$, and since $a^2+a-1 > a^2+2a-2$, this expression does not exceed

$$\left(\frac{1-\eta}{1-a} + o(1)\right) r^{(1-a)/a} + \log \sigma(r) + O(1);$$

consequently, (2.1) is determined if for some $\eta > 0$

$$\mu_n^{1/(2n^a)} = o\left\{\exp\left(\frac{1-\eta}{1-a} n^{1-a}\right)\right\}.$$

(iii) Let $3^{1/2}-1 > a > 0$. If we suppose that

$$(3.8) \quad G(r) \leq Br^{a/2},$$

for some $B > 0$, the right-hand side of (3.5) does not exceed

$$A + Br^{a/2} + C_1 a^{1/2} r^{(a-1)/(2a)} + BC_2 a^{1/2} \rho^{(a-1)/a} r^{(1-a)/a} \\ = r^{(1-a)/a} (BC_2 a^{-1/2} \rho^{(a-1)/a} + o(1)) + O(1).$$

If ρ is so large that $BC_2 a^{1/2} \rho^{(a-1)/a} < 1/(1-a)$, (3.5) is clearly impossible for large r ; consequently, (2.1) is determined if (3.8) is satisfied; that is, if

$$\mu_n^{1/(2n^a)} \leq \exp\{O(n^{a/2})\}.$$

Conditions for the cases $a = 2(2^{1/2}-1)$ or $a = 3^{1/2}-1$ are easily written.

CAMBRIDGE, ENGLAND

THE BOUNDARY PROBLEM OF AN ORDINARY LINEAR DIFFERENTIAL SYSTEM IN THE COMPLEX DOMAIN*

BY
RUDOLPH E. LANGER

1. **Introduction.** The subject of this discussion is to be the system of ordinary linear differential equations which is of the form, or is reducible to the form,

$$(1.1) \quad y_i'(x, \lambda) = \left\{ \lambda r_i(x) + \sum_{v=1}^n q_{i,v}(x, \lambda) \right\} y_v(x, \lambda), \quad i = 1, 2, \dots, n,$$

the variable x and the parameter λ being complex, $|\lambda|$ being indefinitely large, and the coefficients $q_{i,j}(x, \lambda)$ being bounded.† Specifically the matters to be considered are:

In Part I, the dependence of solutions of the system upon λ , when the modulus of the latter is large, and the domain of x is a suitable finite portion of the complex x plane;

In Part II, the boundary problem which arises when a set of conditions applying at any suitable finite set of points of the x domain is imposed upon the system;

In Part III, the theory of the expansibility of a set of n arbitrary analytic functions of x in series of characteristic solutions of the boundary problem.

These matters have, of course, all been widely investigated before this, and discussions of them are to a large extent classical in the literature. However, these discussions—and the present one is, to be sure, no exception in this respect—are invariably restricted in their scope in one way or another by being of necessity based upon hypotheses which to a greater or less extent delimit the considerations and the applicability of the results. Such restrictive hypotheses may, of course, be essential, in the sense that they serve to delimit the considerations to intrinsically identifiable cases of a problem of

* Presented to the Society, April 9, 1938; received by the editors January 9, 1939, and in revised form, February 14, 1939.

† The reduction of differential systems

$$u_i'(x, \lambda) = \sum_{v=1}^n \{ \lambda a_{i,v}(x) + b_{i,v}(x, \lambda) \} u_v(x, \lambda), \quad i = 1, 2, \dots, n,$$

to form (1) is considered in G. D. Birkhoff and R. E. Langer, *The boundary problems and developments associated with a system of ordinary linear differential equations of the first order*, Proceedings of the American Academy of Arts and Sciences, vol. 58 (1923), pp. 72-74.

excessive generality. On the other hand, they may be unessential in the sense that they are primarily called forth by shortcomings of the methods used, or by inadequate or otherwise faulty formulations of the problems themselves. It is believed that the present paper contributes something to a removal or relaxation of several hypotheses of the latter category upon which related earlier discussions are dependent.

The features in which the present paper differs most markedly from previous ones include the following.

(a) With only few and fragmentary exceptions the problems dealt with have heretofore been considered only in the cases of a real variable. The discussion here is allocated to the complex plane, and so includes the earlier results as special cases.

(b) The complete dependence of the functional forms of the solutions of a system of the type (1.1) upon the parameter λ , when $|\lambda|$ is large, has been derived heretofore, even for a real variable, only under the heavily restrictive hypothesis that the coefficient functions $r_i(x)$, as complex quantities, are such that their differences $\{r_i(x) - r_j(x)\}$ all maintain constant arguments over the x range considered. The present discussion is not so restricted, and hence materially extends the existing theory, this being so even when the variable is specialized to be real.

(c) The boundary problems which have been studied in connection with the system (1.1) have almost exclusively been such as arise when the boundary conditions apply only at collinear points, that is, generally points of the axis of reals. In the present paper these conditions are permitted to apply at any finite set of points within appropriate regions of the complex plane. This generalization calls for corresponding generalizations of many familiar notions, among them those of the adjoint boundary problem, of the Green's function, of regularity of the boundary problem, and so on; and such generalizations are given.

(d) Heretofore the theory of the expansibility of an arbitrary vector, that is, of a set of n arbitrary functions, in terms of the characteristic solutions of a regular boundary problem, has been given only for the very restricted cases in which the coefficient functions $r_i(x)$, as complex quantities, each maintain a constant argument for all values of x involved. In the present paper this restriction is dispensed with.

The system (1.1) is notationally treated, in the following, in its matrix form. Insofar as deductions of a formal nature are concerned, those here given include as special cases almost all those which have become classical for the cases of a real variable, whether the boundary conditions are taken to apply at just two, or at more than two, points. In a number of instances the present

formulations are thought to embody material improvements, even when they are specialized to the ranges of the earlier discussions. This seems to be so particularly in the cases of boundary conditions applying at intermediate points of the x interval as well as at the end points. In the rigorous analysis no attempt has been made to pare the hypotheses down to a minimum, or to sharpen the deductions to any point at which they would include any major portion of the many refined and precise results which are known in the case of a real variable. To do that would have extended the bounds of the paper excessively.

PART I. THE FORMS OF THE SOLUTIONS WHEN $|\lambda|$ IS LARGE

2. **The matrix equation.** If $\mathcal{Y}(x, \lambda)$ is a matrix* which satisfies the differential matrix equation

$$(2.1) \quad \mathcal{Y}'(x, \lambda) = \{\lambda \mathcal{R}(x) + \mathcal{Q}(x, \lambda)\} \mathcal{Y}(x, \lambda),$$

in which the prime indicates differentiation with respect to x , and in which the coefficient matrices are

$$(2.2) \quad \mathcal{R}(x) \equiv (\delta_{i,j} r_i(x)), \dagger \quad \mathcal{Q}(x, \lambda) \equiv (q_{i,j}(x, \lambda)),$$

then the elements of any column of $\mathcal{Y}(x, \lambda)$ comprise a solution of the differential system (1.1). If $\mathcal{Y}(x, \lambda)$ is nonsingular, its columns are linearly independent and so yield a complete set of solutions of (1.1). The matrix equation (2.1) may, therefore, be chosen to replace completely the scalar system (1.1) as the basis of the discussion. This will henceforth be done, because of the notational advantages which are thereby to be gained.

The equation (2.1) is to be considered with the parameter λ complex and ranging over some suitable region of the λ plane in which $|\lambda|$ is unbounded. The variable x is likewise to be complex, and is to range either over some suitable bounded region, or over some suitable finite curvilinear arc. The term *suitable* as here used needs, of course, to be made precise. In the case of an x region, that is, of a two-dimensional x domain, this is to be done by the definition:

A pair of regions in the x and λ planes will be said to be suitable to the matrix equation (2.1), if for x and λ within them the coefficients of the equation fulfill the specifications:

(a) *the functions $r_i(x)$, ($i=1, 2, \dots, n$), are analytic and bounded, and their differences $\{r_i(x) - r_j(x)\}$, ($i \neq j$), are all bounded from zero;*

* Throughout the paper square matrices of order n will be designated by means of German capital letters, and these letters will be used solely in this sense. The elements of a matrix will then generally be designated by the corresponding lower case italic letters, that is, in the manner $\mathfrak{a}_{i,j}(x, \lambda)$.

† The symbol $\delta_{i,j}$ will always be used in the sense $\delta_{i,j}=0$, if $i \neq j$; $\delta_{i,i}=1$, if $i=j$.

(b) the functions $q_{i,j}(x, \lambda)$, ($i, j = 1, 2, \dots, n$), are analytic and bounded in x and in λ , and, when $|\lambda|$ is sufficiently large, admit of either actual or asymptotic representations, such that

$$(2.3) \quad \mathfrak{Q}(x, \lambda) \sim \sum_{h=0}^{\infty} \lambda^{-h} \mathfrak{Q}^{(h)}(x),$$

the elements of the matrices $\mathfrak{Q}^{(h)}(x)$ being analytic and bounded.

If the domain of x is one-dimensional, that is, an arc, as it is in the case of a real variable, the definition of the term *suitable* is to be that obtained from the definition above when the term *analytic*, as used relative to x , is replaced by *indefinitely differentiable along the arc*.

To assure the existence of a basis for the entire discussion at hand, this will be assumed as

HYPOTHESIS (i). The given differential matrix equation (2.1) is one for which there exist some suitable regions of the x and λ planes.

If in the equation (2.1) the substitution

$$(2.4) \quad \mathfrak{Y}(x, \lambda) = (\delta_{i,j} e^{\lambda \omega x + \phi_j(x)}) \mathfrak{U}(x, \lambda),$$

is made, with ω any constant and the $\phi_j(x)$, ($j = 1, 2, \dots, n$), any analytic functions, the equation satisfied by the matrix $\mathfrak{U}(x, \lambda)$ is found to be of the same form as (2.1), and to differ from the latter by having the functions $\{r_j(x) - \omega\}$ in the place of the $r_j(x)$, and the functions $\{q_{i,j}(x, \lambda) - \phi_j'(x)\} \exp \{\phi_j(x) - \phi_i(x)\}$ in the place of the $q_{i,j}(x, \lambda)$. From this it may be observed firstly, that since ω can always be chosen so that the determinant of the matrix $(\delta_{i,j} [r_i(x) - \omega])$ is not zero, any given matrix equation (2.1) is transformable into another such equation in which the matrix filling the role of $\mathfrak{R}(x)$ is nonsingular. Secondly, it may be noted that since the functions $\phi_j(x)$ may be chosen so that $\phi_j'(x) \equiv q_{j,j}^{(0)}(x)$, any given equation (2.1) is always transformable into one in which the elements of the main diagonal of the coefficient $\mathfrak{Q}^{(0)}(x)$ of (2.3) all vanish identically. Of these facts, that concerning $\mathfrak{R}(x)$ yields no immediate advantage, though it will later be referred to. That concerning $\mathfrak{Q}(x, \lambda)$ does yield an advantage, and hence it will be assumed forthwith that (2.1) represents such a transformation of the given matrix equation that in it

$$(2.5) \quad q_{i,j}^{(0)}(x) \equiv 0, \quad j = 1, 2, \dots, n.$$

From classical and familiar existence theorems it is known that an equation of the form (2.1) possesses solutions which are nonsingular matrices whose elements are analytic functions of x and λ . Moreover, if any particular

such solution is designated by $\mathcal{Y}^{(p)}(x, \lambda)$, then the general solution of the equation is obtained from the formula

$$(2.6) \quad \mathcal{Y}(x, \lambda) = \mathcal{Y}^{(p)}(x, \lambda)\mathfrak{C},$$

by permitting \mathfrak{C} to represent any matrix whose elements are independent of x . The matrix \mathfrak{C} may, of course, depend upon λ .

3. "Associated" regions and "fundamental" regions. In any given suitable region of the x plane, a set of analytic functions $R_i(x)$ may be chosen such that their derivatives are

$$(3.1) \quad R'_i(x) \equiv r_i(x), \quad i = 1, 2, \dots, n,$$

and these functions will be bounded. We suppose such a choice to have been made. Then if λ is regarded for the moment as fixed, in some suitable λ region, each one of the relations

$$(3.2) \quad \xi^{(i,j)} = \lambda \{ R_i(x) - R_j(x) \}, \quad i, j = 1, 2, \dots, n; i \neq j,$$

defines its left-hand member as a complex variable, and maps any closed sub-region X of the given suitable x region upon a corresponding closed region $\Xi^{(i,j)}$ in the respective $\xi^{(i,j)}$ plane. Consider now the possibility of such a sub-region X containing a set of points $x_*^{(i,j)}$, not necessarily distinct, which have the properties that the point $\xi_*^{(i,j)}$ which lies in the $\xi^{(i,j)}$ plane and corresponds to $x_*^{(i,j)}$ under (3.2), admits of connection with each and every point of the respective region $\Xi^{(i,j)}$ by some curve of bounded length, which lies entirely in $\Xi^{(i,j)}$ and upon which the abscissa is a non-increasing function of the arc length as measured from $\xi_*^{(i,j)}$. This possibility is readily seen to be contingent directly upon the shapes of the regions $\Xi^{(i,j)}$, and hence upon the shape of the region X . When such points $x_*^{(i,j)}$ do exist, they evidently lie upon the boundary of the region X , and the points $\xi_*^{(i,j)}$ are clearly boundary points of maximum abscissa of the respective regions $\Xi^{(i,j)}$.

If λ is now allowed to vary, it is clear at once from (3.2), that all changes in $|\lambda|$ produce in the several $\xi^{(i,j)}$ planes merely changes of scale. Such changes cannot, therefore, influence either the existence or the location of any point $x_*^{(i,j)}$. On the other hand, any change in $\arg \lambda$ produces a rotation of each $\xi^{(i,j)}$ plane, and hence, in particular, of each of the regions $\Xi^{(i,j)}$. Such a rotation may deprive the points of an existing set $x_*^{(i,j)}$ of their characteristic properties. However, it does not necessarily do so, it being possible for a set $x_*^{(i,j)}$ to remain independent of λ and retain its properties under a variation of $\arg \lambda$ over some specific range. This will be shown below. The possibility is again contingent upon the shape of the region X , but also upon the range of $\arg \lambda$ which is in question. We make the definition

A closed subregion X of a suitable x region and a subregion Λ of a suitable λ region will be termed "associated" regions if there exists in X a set of points $x_*^{(i,j)}$, ($i, j = 1, 2, \dots, n; i \neq j$), not necessarily distinct, but fixed as to λ , having the properties described above, and retaining them for all λ in the region Λ .

A given suitable λ region may not admit of any region of the x plane being associated with it. It may, however, still admit of being completely covered by subregions each of which admits of association with some x region. The x regions here in question may, moreover, in some cases have a part in common. We make the definition

A closed region of the x plane will be designated as a fundamental region relative to a given suitable λ region, if it is included in each of a finite number of regions X , which are associated with regions Λ completely covering the suitable λ region in question.

Finally it may be observed that if a given suitable λ region is bounded by lines along which $\arg \lambda$ is constant, that is, if it is a sector (or the part of a sector in which $|\lambda| > N^\dagger$), then since only $\arg \lambda$ comes into question, any subregion Λ which is associated with an x region may also be taken as a sector (or the part of it in which $|\lambda| > N$).

4. The existence of associated and fundamental regions. Inasmuch as regions to be termed associated or fundamental have been defined only in terms of properties prescribed for them, the question of their existence must be considered. In this connection the following will be shown

If x_0 and λ_0 are arbitrarily chosen interior points of a suitable two-dimensional x region and a suitable λ region, respectively, then there exist associated regions of which they are likewise interior points, and there exist x regions which are fundamental relative to the suitable λ region and having x_0 in their interiors.

In the case $n = 2$ these facts are evident almost by inspection. For in this case the only variables defined by (3.2) are $\xi^{(1,2)}$ and $\xi^{(2,1)}$, and these are negatives of each other. Let X , therefore, be taken as any such part of the given suitable x region as contains x_0 in its interior, and as maps in the $\xi^{(1,2)}$ plane, under (3.2), with $\lambda = \lambda_0$, upon a convex polygon with no side parallel to the axis of imaginaries. This polygon is then the region $\Xi^{(1,2)}$, and clearly the region $\Xi^{(2,1)}$ is also such a polygon. The extreme right-hand vertices of these polygons evidently fill, respectively, the specifications upon the points

[†] The symbol $|\lambda| > N$, which there will be frequent occasion to use, is to be read as a mere abbreviation of the phrase "when $|\lambda|$ is sufficiently large." The letter N is, therefore, not to be regarded as designating always one and the same number, but as designating in each case *some* number, possibly different ones in different recurrences of the symbol. The precise magnitude of N is generally left undiscussed as not germane to the argument.

$\xi_*^{(1,2)}$ and $\xi_*^{(2,1)}$. If λ is now allowed to vary, the resulting rotations under which each polygon maintains some one vertex in the extreme right-hand position determine ranges of $\arg \lambda$, and hence subregions of the given λ region, which are associated with the region X . Of these, one contains λ_0 in its interior. Since any given suitable λ region may clearly be covered by a finite number of such subregions, the region X which was chosen is seen to be a fundamental one relative to any suitable portion of the λ plane.

If $n > 2$ the reasoning may be fashioned as follows. With any choice of a real number τ_0 , the interval $(0, \pi)$ is divided into at most $n(n-1)$ subintervals by those of its points which are congruent, modulo π , to the points of the set

$$(4.1) \quad \tau_0^{(i,j)} = \tau_0 + \arg \lambda_0 + \arg \{r_i(x_0) - r_j(x_0)\}, \\ i, j = 1, 2, \dots, n; i \neq j.$$

Of these subintervals at least one is of a length 2δ , with $\delta \geq \pi/2n(n-1)$, and with a proper choice of τ_0 this subinterval is bisected by the point $\pi/2$. We suppose τ_0 so chosen. Then each of the points (4.1) is congruent, modulo π , to some point of the closed interval $(-\pi/2 + \delta, \pi/2 - \delta)$. Let ϵ_1, ϵ_2 , and ϵ_3 be chosen as positive constants subject to the restriction

$$(4.2) \quad \epsilon_1 + \epsilon_2 + \epsilon_3 < \delta,$$

but otherwise arbitrary.

Consider now any curve C in the given x region, which

(a) lies in a neighborhood of the point x_0 in which the relations

$$(4.3) \quad |\arg \{r_i(x) - r_j(x)\} - \arg \{r_i(x_0) - r_j(x_0)\}| < \epsilon_1, \\ i, j = 1, 2, \dots, n; i \neq j,$$

are all fulfilled, and

(b) has a continuously turning tangent whose inclination τ satisfies the condition

$$(4.4) \quad |\tau - \tau_0| < \epsilon_2.$$

Finally let $\arg \lambda$ be restricted by the relation

$$(4.5) \quad |\arg \lambda - \arg \lambda_0| \leq \epsilon_3.$$

For any set of indices (i, j) , the arc C corresponds under (3.2) to an arc $\Gamma^{(i,j)}$ in the plane of the variable $\xi^{(i,j)}$. If the inclination of the tangent line to $\Gamma^{(i,j)}$ is denoted by $\tau^{(i,j)}$, it follows from (3.2) that $\tau^{(i,j)} = \tau + \arg \lambda + \arg \{r_i(x) - r_j(x)\}$, and hence, from (4.4), (4.5), (4.3), (4.2), that $|\tau^{(i,j)} - \tau_0^{(i,j)}| < \delta$. Thus $\tau^{(i,j)}$ is bounded from becoming congruent, modulo π , with either of the values $-\pi/2$ or $\pi/2$; that is, the slope of $\Gamma^{(i,j)}$ is bounded.

Let X be chosen now as any subregion of the given suitable x region which contains x_0 in its interior, and which is bounded by a pair of arcs of the type C described. The corresponding region $\Xi^{(i,j)}$, for each (i, j) , is then bounded by a pair of arcs of the type $\Gamma^{(i,j)}$. These arcs intersect, and one of their intersections, the extreme right-hand point of $\Xi^{(i,j)}$, fills the specifications on the point $\xi_*^{(i,j)}$. It does this, moreover, for all values of λ which are admitted by (4.5). Thus (4.5) determines a λ subregion which is associated with the region X chosen.

Since the constant ϵ_3 is not dependent upon λ_0 , it is clear that in any suitable λ region a finite number of points may be chosen so that they fill the role of λ_0 above, and such that the entire λ region is covered by the corresponding subregions (4.5). Each of these subregions, it has been shown, is associated with some region X which contains x_0 in its interior. The part common to these regions X is seen at once to be a fundamental region relative to the given suitable λ region.

The discussion thus given was based explicitly upon the assumption that the suitable x region containing x_0 was a two-dimensional one. If the x region is one-dimensional, that is, an arc, the discussion is not generally applicable, and no association of x and λ regions may be possible. Exceptional in this respect is the case in which some segment of the x domain maps under each of the transformations (3.2) upon a straight segment, that is, if on such a segment

$$(4.6) \quad \arg \{R_i(x) - R_j(x)\} \equiv \text{constant}, \quad i, j = 1, 2, \dots, n; i \neq j.$$

In this case the argument given above serves without modification to show that the x segment in question is a fundamental region relative to any suitable λ region.

The conditions (4.6) will be recognized as an important part of the hypotheses upon which the discussions analogous to that of this paper, but applying to the real variable x , have classically been based. The motivation for this is thus seen to lie in the need of having the basic domain of the variable be a fundamental region.

5. The solution of an approximating equation. If x and λ are taken in any suitable regions, and the matrices $\mathfrak{Q}^{(h)}(x)$ are those of (2.3), the formulas

$$(5.1) \quad \begin{aligned} p_{i,j}^{(0)} &\equiv \delta_{i,j}, & i, j &= 1, 2, \dots, n; \\ p_{i,j}^{(l)}(x) &\equiv \{r_i(x) - r_j(x)\}^{-1} \left\{ p_{i,j}^{(l-1)'}(x) - \sum_{h=0}^{l-1} \sum_{v=1}^n q_{i,v}^{(l-h)}(x) p_{v,j}^{(h)}(x) \right\}, & i &\neq j; \\ p_{i,i}^{(l)'}(x) &\equiv \sum_{h=0}^l \sum_{v=1}^n q_{i,v}^{(l-h)}(x) p_{v,i}^{(h)}(x), \end{aligned}$$

together with any choice of constants of integration, define in succession for $l=0, 1, 2, 3$, and so on, a sequence of matrices $\mathfrak{P}^{(l)}(x)$. These matrices are analytic and bounded, and satisfy the matrix equations

$$(5.2) \quad \mathfrak{P}^{(l+1)}(x)\mathfrak{R}(x) - \mathfrak{R}(x)\mathfrak{P}^{(l+1)}(x) + \mathfrak{P}^{(l)'}(x) - \sum_{h=0}^l \mathfrak{Q}^{(l-h)}(x)\mathfrak{P}^{(h)}(x) = \mathfrak{O}, \\ l = 0, 1, 2, \dots$$

Let any natural number k be chosen, then, and let the formulas

$$(5.3) \quad \mathfrak{P}_k(x, \lambda) \equiv \sum_{l=0}^k \lambda^{-l} \mathfrak{P}^{(l)}(x),$$

$$(5.4) \quad \mathfrak{E}(x, \lambda) \equiv (\delta_{ij} e^{\lambda R_i(x)}),$$

define their left-hand members. The functions $R_i(x)$ are those of (3.2). Then in virtue of the relations (2.3) and (5.2), it is found directly that the matrix

$$(5.5) \quad \mathfrak{S}(x, \lambda, k) \equiv \mathfrak{P}_k(x, \lambda) \mathfrak{E}(x, \lambda),$$

is such that

$$\mathfrak{S}' - \{\lambda \mathfrak{R} + \mathfrak{Q}\} \mathfrak{S} \sim \left\{ \lambda^{-k} \mathfrak{P}^{(k)'} - \sum_{h=k}^{\infty} \lambda^{-h} \sum_{l=0}^k \mathfrak{Q}^{(h-l)} \mathfrak{P}^{(l)} \right\} \mathfrak{E}.$$

When $|\lambda| > N$, this can be written, since the matrix (5.3) is then certainly nonsingular, in the form

$$(5.6) \quad \mathfrak{S}' = \{\lambda \mathfrak{R}(x) + \mathfrak{Q}(x, \lambda) + \lambda^{-k} \mathfrak{A}(x, \lambda, k)\} \mathfrak{S},$$

in which the coefficient matrix $\mathfrak{A}(x, \lambda, k)$ is one which admits of a representation

$$(5.7) \quad \mathfrak{A}(x, \lambda, k) \sim \left\{ \mathfrak{P}^{(k)'}(x) - \sum_{h=k}^{\infty} \lambda^{-h+k} \sum_{l=0}^k \mathfrak{Q}^{(h-l)}(x) \mathfrak{P}^{(l)}(x) \right\} \mathfrak{P}_k^{-1}(x, \lambda).$$

The equation (5.6) is a matrix differential equation which in an obvious sense approximates the given equation (2.1) when $|\lambda|$ is large. The matrix (5.5) is thus seen to be a nonsingular analytic solution of an equation which approximates the given equation when $|\lambda| > N$.

Let $\mathfrak{Y}^{(p)}(x, \lambda)$ designate any particular nonsingular analytic solution of the equation (2.1). Then in virtue of the equation (5.6) the relation

$$(5.8) \quad \{\mathfrak{S}^{-1} \mathfrak{Y}^{(p)}\}' = -\lambda^{-k} \mathfrak{S}^{-1} \mathfrak{A} \mathfrak{Y}^{(p)}$$

is readily found to be an identity. In terms of the matrices defined by the formulas

$$(5.9) \quad \mathfrak{Y}_{h,l} = (\delta_{ij} \delta_{l,i}), \quad h, l = 1, 2, \dots, n,$$

in each of which one element is unity and all others are zero, the relation (5.8) may be written in the alternative form

$$(5.10) \quad \{\mathfrak{S}^{-1}\mathfrak{Y}^{(p)}\}' + \lambda^{-k} \sum_{h,l=1}^n \mathfrak{Z}_{h,h} \mathfrak{S}^{-1}\mathfrak{Y}^{(p)} \mathfrak{Z}_{l,l} = \mathfrak{D}.$$

This form is convenient for the use to which the relation is to be put below.

6. The solutions of the given equation. For the formal deductions of the preceding section it sufficed to regard x and λ as in any suitable regions. Let them be restricted now to any pair of associated regions X and Λ . There exists then in X a set of points $x_{*}^{(h,l)}$ as described in §3, and these points do not depend upon λ . By an appropriate integration based upon these points, the relation (5.10) may be given the form

$$(6.1) \quad \mathfrak{S}^{-1}(x) \mathfrak{Y}^{(p)}(x) + \lambda^{-k} \sum_{h,l=1}^n \int_{x_{*}^{(h,l)}}^x \mathfrak{Z}_{h,h} \mathfrak{S}^{-1}(x_1) \mathfrak{Y}(x_1) \mathfrak{Y}^{(p)}(x_1) \mathfrak{Z}_{l,l} dx_1 = \mathfrak{R}(\lambda),$$

and this may be looked upon as defining its right-hand member as an analytic matrix independent of x .

Let $\mathfrak{C}(\lambda)$ be a matrix which is unspecified, except that $\mathfrak{C}(\lambda) \neq \mathfrak{D}$, and let

$$(6.2) \quad \mathfrak{B}(x, \lambda) \equiv \mathfrak{Y}^{(p)}(x) \mathfrak{C}(\lambda) \mathfrak{S}^{-1}(x).$$

If it is observed that by (5.9)

$$\sum_{l=1}^n \mathfrak{Z}_{l,l} \mathfrak{C}(\lambda) = \sum_{l=1}^n \mathfrak{C}(\lambda) \mathfrak{Z}_{l,l},$$

it is found that on multiplication by $\mathfrak{S}(x)$ on the left, and by $\mathfrak{C}(\lambda) \mathfrak{S}^{-1}(x)$ on the right, the relation (6.1) becomes

$$(6.3) \quad \mathfrak{B}(x) + \lambda^{-k} \sum_{h,l=1}^n \mathfrak{S}(x) \mathfrak{Z}_{h,h} \mathfrak{S}^{-1}(x_1) \mathfrak{Y}(x_1) \mathfrak{B}(x_1) \mathfrak{S}(x_1) \mathfrak{Z}_{l,l} \mathfrak{S}^{-1}(x) dx_1 \\ = \mathfrak{S}(x) \mathfrak{R}(\lambda) \mathfrak{C}(\lambda) \mathfrak{S}^{-1}(x).$$

Now from (5.5) and (5.9) it is seen that

$$\mathfrak{S}(x) \mathfrak{Z}_{h,h} \mathfrak{S}^{-1}(x_1) \equiv \mathfrak{P}_h(x) \mathfrak{Z}_{h,h} \mathfrak{P}_h^{-1}(x_1) \exp [\lambda \{R_h(x) - R_h(x_1)\}],$$

and

$$\mathfrak{B}(x_1) = \sum_{\alpha,\beta=1}^n \mathfrak{Z}_{\alpha,\beta} \mathfrak{V}_{\alpha,\beta}(x_1).$$

If $\xi_1^{(i,j)}$ is the value which corresponds under (3.2) to x_1 , it follows that (6.3) finally takes the form

$$(6.4) \quad \mathfrak{B}(x, \lambda) + \lambda^{-k} \mathfrak{B}(x, \lambda) = \mathfrak{S}(x) \mathfrak{R}(\lambda) \mathfrak{C}(\lambda) \mathfrak{S}^{-1}(x),$$

in which

$$(6.5) \quad \mathfrak{F}(x, \lambda) = \sum_{h, l, \alpha, \beta=1}^n \int_{x_*(h, l)}^x \mathfrak{P}_k(x) \mathfrak{F}_{h, h} \mathfrak{P}_k^{-1}(x_1) \mathfrak{A}(x_1) \mathfrak{F}_{\alpha, \beta} \mathfrak{P}_k(x_1) \mathfrak{F}_{l, l} \mathfrak{P}_k^{-1}(x) \\ \cdot e^{\xi(h, l) - \xi_1(h, l)} v_{\alpha, \beta}(x_1) dx_1.$$

The elements of the matrix (6.2) are analytic in X , and this region, is by definition, closed. Its elements, therefore, take on numerical maxima; hence there exists a scalar $m(\lambda)$ independent of x , such that

$$(6.6) \quad |v_{i, j}(x)| \leq m(\lambda), \quad i, j = 1, 2, \dots, n,$$

and the equality holds for some index pair (i, j) , at some point x . Moreover $m(\lambda) > 0$, since by hypothesis $\mathfrak{C}(\lambda) \neq \mathfrak{O}$.

Let the path of integration from the point $x_*(h, l)$ to x in (6.5) be taken now as a curve along which the real part of $\xi_1(h, l)$ is non-increasing. It is precisely the characteristic property of the point $x_*(h, l)$ that, whatever x may be, there exists such a path. During the integration, then, it is clear that

$$|e^{\xi(h, l) - \xi_1(h, l)}| \leq 1.$$

Finally since the matrices in the integrands of (6.5) are obviously bounded, both as to x and as to λ , when $|\lambda| > N$, there exists an absolute scalar constant M such that the elements of (6.5) satisfy the relation

$$(6.7) \quad |h_{i, j}(x, \lambda)| \leq Mm(\lambda), \quad i, j = 1, 2, \dots, n.$$

For that pair of indices, and that point x , for which the equality holds in (6.6), therefore,

$$(6.8) \quad m(\lambda) \left\{ 1 - \frac{M}{|\lambda|^k} \right\} \leq \left| v_{i, j}(x, \lambda) + \frac{h_{i, j}(x, \lambda)}{\lambda^k} \right|.$$

Since the left-hand member of this is positive, when $|\lambda| > N$, it follows that the matrix on the left of (6.4) is not the zero matrix. This must, therefore, be so for the matrix on the right, namely, $\mathfrak{R}(\lambda)\mathfrak{C}(\lambda) \neq \mathfrak{O}$. Since this follows with $\mathfrak{C}(\lambda)$ unspecified except for $\mathfrak{C}(\lambda) \neq \mathfrak{O}$, it must be concluded that the matrix $\mathfrak{R}(\lambda)$ is nonsingular.

With the existence of the matrix $\mathfrak{R}^{-1}(\lambda)$ thus established, we may now choose $\mathfrak{C}(\lambda) = \mathfrak{R}^{-1}(\lambda)$. The right-hand member of (6.4), and hence the left-hand member, reduces then to the unit matrix. Since the right-hand member of (6.8) is thus at most unity, it follows that the function $m(\lambda)$ is bounded, when $|\lambda| > N$. Then by (6.7) the elements of the matrix $\mathfrak{F}(x, \lambda)$ are bounded. From (6.4) and (6.2), lastly,

$$(6.9) \quad \mathfrak{Y}^{(p)}(x, \lambda)\mathfrak{R}^{-1}(\lambda) = \{\mathfrak{F} - \lambda^{-k}\mathfrak{F}(x, \lambda)\}\mathfrak{C}(x, \lambda, k),$$

and in this the left-hand member is a nonsingular analytic solution of the given equation (2.1). The existence of a solution of this form (6.9) is what was to be established. The result may be formulated thus:

If x and λ are restricted to any pair of associated regions, there exists an analytic solution of the equation (2.1) which is of the form

$$(6.10) \quad \mathcal{Y}(x, \lambda) = \mathcal{P}(x, \lambda)\mathcal{E}(x, \lambda),$$

the matrix $\mathcal{P}(x, \lambda)$ being of the form

$$(6.11) \quad \mathcal{P}(x, \lambda) = \mathcal{I} + \sum_{h=1}^{k-1} \lambda^{-h} \mathcal{P}^{(h)}(x) + \lambda^{-k} \mathcal{B}_k(x, \lambda).$$

In this k may be taken as any natural number, and the elements of the matrix $\mathcal{B}_k(x, \lambda)$ are bounded when $|\lambda| > N$.

This result may be extended at once to the case in which λ is restricted merely to some suitable region, while x , on the other hand, remains in a region which is fundamental relative to the λ region in question. In this case the λ region may be subdivided into a finite number of subregions, in each of which some solution of the equation maintains the form (6.10), (6.11). The solutions which respectively maintain these forms in different λ subregions will in general be different. In virtue of (2.6) it is clear that in any λ subregion the general solution of the equation (2.1) is of the form

$$\mathcal{Y}(x, \lambda) = \mathcal{P}(x, \lambda)\mathcal{E}(x, \lambda)\mathcal{C}(\lambda),$$

with $\mathcal{P}(x, \lambda)$ as given by (6.11).

PART II. THE BOUNDARY PROBLEM

7. Definition and qualitative aspects of the problem. If any finite set of points $\eta_1, \eta_2, \dots, \eta_m$ with $m \geq 2$ is chosen in any x region suitable to the differential system (1), and the variable is restricted to this domain, while the parameter is restricted to some suitable λ region, the solution of the differential system may be conditioned relative to this set of points by a set of relations

$$\sum_{\mu=1}^m \sum_{\nu=1}^n w_{i,\nu}^{(\mu)} y_{\nu}(\eta_{\mu}, \lambda) = 0, \quad i = 1, 2, \dots, n.$$

Such relations are then termed boundary conditions, and the differential system together with such boundary conditions is said to constitute a boundary problem. The coefficients $w_{i,j}^{(\mu)}$ involved in the boundary conditions may be constants, or may more generally depend analytically upon the parameter λ . They are, of course, independent of x .

The functions which together make up any solution of the differential system are, as has been seen, n in number, constituting the elements of a column of some matrix solution of the differential equation (2.1). They may, therefore, be considered as a vector $\eta(x)$.† If they satisfy the boundary conditions, this vector then satisfies the relations

$$(7.1a) \quad \mathcal{Y}'(x, \lambda) = \{\lambda \mathcal{R}(x) + \mathcal{Q}(x, \lambda)\} \mathcal{Y}(x, \lambda),$$

$$(7.1b) \quad \sum_{\mu=1}^m \mathcal{B}^{(\mu)}(\lambda) \mathcal{Y}(\eta_{\mu}, \lambda) = \mathcal{D}.$$

The boundary problem is thus formulated as that of finding a vector solution of the problem (7.1).

In §2 it was observed that the general solution of the matrix equation (2.1), that is, of (7.1a), is expressible in terms of any particular nonsingular analytic solution $\mathcal{Y}(x, \lambda)$ by means of the formula (2.6). A solution is, therefore, a vector if and only if $\mathcal{C}(\lambda)$ is a vector, and the general vector solution of (7.1a) is thus given by the formula

$$(7.2) \quad \eta(x, \lambda) = \mathcal{Y}(x, \lambda) \mathcal{C}(\lambda),$$

the vector $\mathcal{C}(\lambda)$ being arbitrary. If such a vector is to satisfy the relation (7.1b), it follows that the equation

$$(7.3) \quad \mathcal{D}(\lambda) \mathcal{C}(\lambda) = 0,$$

with

$$(7.4) \quad \mathcal{D}(\lambda) = \sum_{\mu=1}^m \mathcal{B}^{(\mu)}(\lambda) \mathcal{Y}(\eta_{\mu}, \lambda),$$

must be fulfilled. Now the solution (7.2) is evidently trivial, that is, $\eta(x) = 0$, if the vector $\mathcal{C}(\lambda)$ is trivial, that is, $\mathcal{C}(\lambda) = 0$. Hence a necessary and sufficient condition for a non-trivial solution of the boundary problem, is the existence of a non-trivial vector $\mathcal{C}(\lambda)$, which satisfies the equation (7.3). Such familiarly exists if and only if the matrix $\mathcal{D}(\lambda)$ is singular, namely if

$$(7.5) \quad D(\lambda) = 0,$$

where $D(\lambda)$ designates the determinant of the matrix (7.4).

† The use of lower case German letters will be reserved to the designation of vectors of n elements or components, and such vectors are to be regarded freely, as may be convenient, as matrices of one row and n columns, or vice versa. This will lead to no ambiguity if it is agreed, and it shall hereby be so agreed, that all multiplications between matrices and vectors, or of vectors by vectors, is to be understood as being in the matrix sense. A vector is, therefore, to be regarded as a matrix of one row and n columns whenever it appears as a left-hand factor, and as a matrix of n rows and one column if it appears as a right-hand factor.

If at any specified value of λ the condition (7.5) is not fulfilled, the boundary problem admits of no solution and is therefore said to be incompatible. On the other hand, if for a specified λ the condition (7.5) is satisfied, that is, if λ is a root of the equation (7.5), the equation (7.3) does admit a non-trivial solution. More explicitly, if at this λ the rank of the determinant $D(\lambda)$ is $(n-r)$, the equation (7.3) is satisfied by r distinct vectors $c(\lambda)$, and these lead through (7.2) to precisely r linearly independent solutions of the problem. The latter is therefore said in this case to be *compatible to the order r* , the term *simply compatible* being used interchangeably with compatible to the order 1. The roots of the equation (7.5), which thus appear as the λ values for which the boundary problem is solvable, are known as *characteristic values*, and the non-trivial solutions of the problem which exist at these values are called *characteristic solutions*. A characteristic value at which the problem is compatible to the order r is said to be of the *index r* . On the other hand, a characteristic value will be said to be of the *multiplicity s* if it is an s -fold zero of the determinant $D(\lambda)$. It will be seen below that the index of a characteristic value cannot exceed its multiplicity.

It must be observed that although the characteristic values, which are intrinsic to the boundary problem, are obtained from the determinant $D(\lambda)$, neither this determinant, nor the corresponding matrix $\mathfrak{D}(\lambda)$, is uniquely determined by the boundary problem. This is due in part to the fact that the solution $\mathfrak{Y}(x, \lambda)$ of the equation (7.1a) to be used in (7.4) was specified only to the extent that it be analytic and nonsingular, and also in part to the fact that the content of the equation (7.1b) is unchanged if it is multiplied on the left by any analytic nonsingular matrix $\mathfrak{C}_1(\lambda)$. Since when $\mathfrak{Y}(x, \lambda)$ is any eligible solution, all such solutions are expressible by (2.6) in the form $\mathfrak{Y}(x, \lambda)\mathfrak{C}_2(\lambda)$, with the matrix $\mathfrak{C}_2(\lambda)$ analytic and nonsingular, it is clear that the role of the matrix $\mathfrak{D}(\lambda)$ may be given at will to any matrix of the form

$$(7.6) \quad \mathfrak{C}_1(\lambda)\mathfrak{D}(\lambda)\mathfrak{C}_2(\lambda),$$

with \mathfrak{C}_1 and \mathfrak{C}_2 analytic and nonsingular. Conversely, of course, the matrix (7.6) is the most general by which $\mathfrak{D}(\lambda)$ may be replaced. Since the determinant of the matrix (7.6) differs from $D(\lambda)$ only by nonvanishing factors, it is clear that (7.5) as an equation is invariant.

The inverse of the matrix $\mathfrak{D}(\lambda)$ is familiarly given by the formula

$$(7.7) \quad \mathfrak{D}^{-1}(\lambda) \equiv \left(\frac{D_{i,j}(\lambda)}{D(\lambda)} \right),$$

with $D_{i,j}(\lambda)$ denoting the cofactor of the element in the i th row and j th column of $D(\lambda)$. It is clear from this that the elements of $\mathfrak{D}^{-1}(\lambda)$ are analytic

except possibly at the characteristic values, where they may have poles.

8. **The adjoint boundary problem.** Let η_0 be chosen arbitrarily as a point of the x region, either distinct from the points $\eta_1, \eta_2, \dots, \eta_m$, or coincident with any one of them. Then with the various matrices involved identified as those which are similarly denoted in (7.1), and with any specific value of λ , there may or may not exist a *parametric matrix* $\mathfrak{A}(\lambda)$ which is independent of x , and is such that the system of relations

$$(8.1a) \quad \mathfrak{Z}^{(h)'}(x, \lambda) = -\mathfrak{Z}^{(h)}(x, \lambda) \{ \lambda \mathfrak{A}(x) + \mathfrak{Q}(x, \lambda) \},$$

$$\mathfrak{Z}^{(h)}(\eta_h, \lambda) + \mathfrak{A}(\lambda) \mathfrak{B}^{(h)}(\lambda) = \mathfrak{D}, \quad h = 1, 2, \dots, m,$$

$$(8.1b) \quad \sum_{\mu=1}^m \mathfrak{Z}^{(\mu)}(\eta_0, \lambda) = \mathfrak{D},$$

admits of solution by a set of m matrices $\mathfrak{Z}^{(h)}(x, \lambda)$, ($h=1, 2, \dots, m$). It is immediately evident that with $\mathfrak{A}(\lambda) = \mathfrak{D}$, the system is uniquely solved, irrespective of the value of λ , by $\mathfrak{Z}^{(h)}(x, \lambda) \equiv \mathfrak{D}$, ($h=1, 2, \dots, m$), and conversely. This solution is trivial. It is, therefore, requisite for a non-trivial solution that $\mathfrak{A}(\lambda) \neq \mathfrak{D}$, and this will accordingly be generally assumed henceforth. If the parametric matrix $\mathfrak{A}(\lambda)$ is one having a single row, that is, is a vector, any eventual solution of the system will obviously also consist of a set of vectors, and vice versa a vector solution can exist only in connection with a parametric vector. Such a solution of the system (8.1) by vectors is the matter of immediate interest to the discussion, and this problem will be referred to henceforth as the boundary problem *adjoint to the problem* (7.1).

The differential matrix equation (8.1a) is familiarly solved by the inverse of any nonsingular solution of the equation (2.1), and its general solution is therefore given by $\mathfrak{C}(\lambda) \mathfrak{Y}^{-1}(x, \lambda)$, in which $\mathfrak{C}(\lambda)$ is arbitrary and $\mathfrak{Y}(x, \lambda)$ may, in particular, be understood to be that analytic solution of (2.1), that is, of (7.1a), which was used in the deductions of the preceding section. Any vector solutions of the equations (8.1a) are, therefore, of the form

$$(8.2) \quad \mathfrak{z}^{(h)}(x, \lambda) = \mathfrak{c}^{(h)}(\lambda) \mathfrak{Y}^{-1}(x, \lambda), \quad h = 1, 2, \dots, m.$$

With these the relations (8.1b) become

$$(8.3) \quad -\mathfrak{c}^{(h)}(\lambda) = \mathfrak{a}(\lambda) \mathfrak{B}^{(h)}(\lambda) \mathfrak{Y}(\eta_h, \lambda), \quad h = 1, 2, \dots, m,$$

$$\sum_{\mu=1}^m \mathfrak{c}^{(\mu)}(\lambda) = \mathfrak{o},$$

and are to be solved by choice of the vectors $\mathfrak{c}^{(h)}(\lambda)$. If these relations are summed, and the formula (7.4) is recalled, the result is found to be

$$(8.4) \quad \mathfrak{a}(\lambda) \mathfrak{D}(\lambda) = \mathfrak{o}.$$

A solution of the adjoint boundary problem can exist, therefore, only in connection with a parametric vector $a(\lambda)$ which satisfies the equation (8.4). A necessary condition for this, since $a(\lambda)$ must differ from 0, is that the matrix $\mathfrak{D}(\lambda)$ be singular, that is, that λ be a root of the equation (7.5). Since the roots of (7.5) are the characteristic values of the boundary problem (7.1), it follows that whenever the latter is incompatible the adjoint problem is insolvable, that is, is likewise incompatible. Conversely, if $\mathfrak{D}(\lambda)$ is singular, and is, say, of the rank $(n-r)$, the equation (8.4) is solvable and determines precisely r distinct parametric vectors $a(\lambda)$. Each of these leads through the formulas (8.3) to a set of vectors $c^{(h)}(\lambda)$, and by (8.2) r linearly independent solutions of the problem (8.1) are then determined. The adjoint problem is thus appropriately described as compatible to the order r . Since in this case λ is a characteristic value for which the problem (7.1) is compatible to precisely the order r , the result may be formulated thus:

A boundary problem and its adjoint problem have the same characteristic values, and at any characteristic value are compatible to the same order.

The boundary conditions (7.1b) and (8.1b) are so interrelated that if $u(x, \lambda)$ is any matrix satisfying the former, and the matrices $\mathfrak{B}^{(h)}(x, \lambda)$, ($h=1, 2, \dots, m$), together with a parametric matrix $\mathfrak{A}(\lambda)$ satisfy the latter, then

$$(8.5) \quad \sum_{\mu=1}^m \int_{\eta_0}^{\eta_\mu} d\{\mathfrak{B}^{(\mu)}(x, \lambda)u(x, \lambda)\} = \mathfrak{D}.$$

In the classical case of a real variable x and two-point boundary conditions, this will be recognized as a familiar relation; indeed one upon which the definition of the adjoint problem is sometimes based. After performance of the integrations, the left-hand member of the relation takes the form

$$\sum_{\mu=1}^m \mathfrak{B}^{(\mu)}(\eta_\mu, \lambda)u(\eta_\mu, \lambda) - \sum_{\mu=1}^m \mathfrak{B}^{(\mu)}(\eta_0, \lambda)u(\eta_0, \lambda).$$

The second of these sums vanishes by (8.1b), while the first may be written

$$- \mathfrak{A}(\lambda) \sum_{\mu=1}^m \mathfrak{B}^{(\mu)}(\lambda)u(\eta_\mu, \lambda).$$

This vanishes by (7.1b).

An alternative definition of the problem adjoint to (7.1), and one which avoids the introduction of the parametric matrix, may be given by choosing the point η_0 in coincidence with one of the points of the set $\eta_1, \eta_2, \dots, \eta_m$, say with η_r . This is the following:

$$\begin{aligned}
 \mathfrak{Z}_1^{(h)'}(x, \lambda) &= -\mathfrak{Z}_1^{(h)}(x, \lambda) \{ \lambda \mathfrak{R}(x) + \mathfrak{Q}(x, \lambda) \}, \quad h = 1, 2, \dots, m, \\
 \mathfrak{Z}_1^{(h)}(\eta_h, \lambda) + \sum_{\mu=1}^m \mathfrak{Z}_1^{(\mu)}(\eta_r, \lambda) \mathfrak{B}^{(h)}(\lambda) &= \mathfrak{D}, \quad h \neq r, \\
 \mathfrak{Z}_1^{(r)}(\eta_r, \lambda) + \sum_{\mu=1}^m \mathfrak{Z}_1^{(\mu)}(\eta_r, \lambda) \{ \mathfrak{B}^{(r)}(\lambda) - \mathfrak{I} \} &= \mathfrak{D},
 \end{aligned}
 \tag{8.6}$$

the equations to be solved by a set of vectors $\mathfrak{z}_1^{(h)}(x, \lambda)$, ($h=1, 2, \dots, m$), which are not all identically zero. The problem in this formulation is amenable to precisely the same deductions and conclusions as were drawn above from the form (8.1). The passage from the one formulation to the other is easily made by means of the relations

$$\begin{aligned}
 \alpha(\lambda) &= \sum_{\mu=1}^m \mathfrak{z}_1^{(\mu)}(\eta_r, \lambda), \quad \mathfrak{z}^{(h)}(x, \lambda) \equiv \mathfrak{z}_1^{(h)}(x, \lambda), \quad h \neq r, \\
 \mathfrak{z}^{(r)}(x, \lambda) &\equiv - \sum_{\mu \neq r} \mathfrak{z}_1^{(\mu)}(x, \lambda).
 \end{aligned}$$

The form (8.1) was preferred above because of its greater symmetry.

9. The Green's matrices. If $f(x, \lambda)$ is any vector that is analytic in the chosen suitable x and λ regions, the equations

$$(9.1a) \quad u'(x, \lambda) = \{ \lambda \mathfrak{R}(x) + \mathfrak{Q}(x, \lambda) \} u(x, \lambda) + f(x, \lambda),$$

$$(9.1b) \quad \sum_{\mu=1}^m \mathfrak{B}^{(\mu)}(\lambda) u(\eta_\mu, \lambda) = 0,$$

define a vector boundary problem which is related to the problem (7.1), being evidently a nonhomogeneous generalization of it. The solution of this problem is expressible in terms of any nonsingular analytic solution of the matrix differential equation (2.1), and may be deduced as follows.

It is verifiable by actual substitution that the formula

$$(9.2) \quad u^{(p)}(x, \lambda) = \int_{\eta_0}^x \mathfrak{Y}(x, \lambda) \mathfrak{Y}^{-1}(x_1, \lambda) f(x_1, \lambda) dx_1,$$

yields a particular solution of the vector differential equation (9.1a). The general solution of this equation is, therefore, given by

$$(9.3) \quad u(x, \lambda) = u^{(p)}(x, \lambda) + \mathfrak{Y}(x, \lambda) c(\lambda),$$

with the vector $c(\lambda)$ arbitrary. With this evaluation, and in virtue of the formula (7.4), the relation (9.1b) becomes

$$(9.4) \quad \sum_{\mu=1}^m \mathfrak{B}^{(\mu)}(\lambda) u^{(p)}(\eta_\mu, \lambda) + \mathfrak{D}(\lambda) c(\lambda) = 0.$$

Thus the solvability of the problem (9.1) depends upon the possibility of a choice of the vector $c(\lambda)$ to satisfy the equation (9.4). Such a choice is evidently possible and unique provided the matrix $\mathfrak{D}(\lambda)$ is nonsingular; that is, provided λ is not a characteristic value. If this is so, the vector $c(\lambda)$ determined by (9.4) yields through (9.3) the solution of the problem (9.1). The result may be explicitly written

$$(9.5) \quad u(x, \lambda) = \int_{\eta_0}^x \mathfrak{Y}(x, \lambda) \mathfrak{Y}^{-1}(x_1, \lambda) f(x_1, \lambda) dx_1 \\ - \sum_{\mu=1}^m \int_{\eta_0}^{\eta_\mu} \mathfrak{G}^{(\mu)}(x, x_1, \lambda) f(x_1, \lambda) dx_1,$$

with

$$(9.6) \quad \mathfrak{G}^{(h)}(x, x_1, \lambda) \equiv \mathfrak{Y}(x, \lambda) \mathfrak{D}^{-1}(\lambda) \mathfrak{B}^{(h)}(\lambda) \mathfrak{Y}(\eta_h, \lambda) \mathfrak{Y}^{-1}(x_1, \lambda), \\ h = 1, 2, \dots, m.$$

The matrices $\mathfrak{G}^{(h)}(x, x_1, \lambda)$, for which the formulas (9.6) are definitive, thus serve for the solution of the problem (9.1) independently of the vector $f(x, \lambda)$ which may be involved therein. They are to be known henceforth as the *Green's matrices*, and are best regarded as associated with the boundary problem (7.1), since they are constructed solely from matrices involved in the latter. They exist and are evidently analytic whenever the associated boundary problem is incompatible. Moreover, they are unique, for though $\mathfrak{Y}(x, \lambda)$ and the $\mathfrak{B}^{(\mu)}(\lambda)$ may at will be replaced by $\mathfrak{Y}(x, \lambda) \mathfrak{C}_2(\lambda)$ and $\mathfrak{C}_1(\lambda) \mathfrak{B}^{(\mu)}(\lambda)$, respectively, as was observed in §7, such a change would call for the replacement of $\mathfrak{D}(\lambda)$ by the matrix (7.6). The formulas (9.6) are evidently invariant under such substitutions.

For subsequent use it may be recorded that the Green's matrices satisfy the relations

$$(9.7a) \quad \sum_{\mu=1}^m \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \equiv \mathfrak{Y}(x, \lambda) \mathfrak{Y}^{-1}(x_1, \lambda),$$

$$(9.7b) \quad \mathfrak{G}^{(h)}(x, \eta_h, \lambda) \equiv \mathfrak{Y}(x, \lambda) \mathfrak{D}^{-1}(\lambda) \mathfrak{B}^{(h)}(\lambda),$$

$$(9.7c) \quad \sum_{\mu=1}^m \mathfrak{B}^{(\mu)}(\lambda) \mathfrak{G}^{(h)}(\eta_\mu, x_1, \lambda) \equiv \mathfrak{B}^{(h)}(\lambda) \mathfrak{Y}(\eta_h, \lambda) \mathfrak{Y}^{-1}(x_1, \lambda), \\ h = 1, 2, \dots, m.$$

These are readily deduced directly from the formulas (9.6). The formula (9.7a) makes possible the reduction of (9.5) to an alternative and more compact form. Since the integrands involved in (9.5) are all analytic, the paths of integration may be chosen at pleasure, and hence may in particular be

chosen to lie in coincidence from the point η_0 to the point x . With this choice, and in virtue of (9.7a), the formula reduces to

$$(9.8) \quad u(x, \lambda) = \sum_{\mu=1}^m \int_{\eta_\mu}^x \mathfrak{G}^{(\mu)}(x, x_1, \lambda) f(x_1, \lambda) dx_1.$$

The nonhomogeneous vector boundary problem which generalizes the adjoint problem (8.1) in a manner similar to the above, is evidently given by the equations

$$(9.9a) \quad v^{(h)'}(x, \lambda) = -v^{(h)}(x, \lambda) \{ \lambda \mathfrak{R}(x) + \mathfrak{Q}(x, \lambda) \} + f(x, \lambda),$$

$$v^{(h)}(\eta_h, \lambda) + a_1(\lambda) \mathfrak{B}^{(h)}(\lambda) = 0, \quad h = 1, 2, \dots, m,$$

$$(9.9b) \quad \sum_{\mu=1}^m v^{(\mu)}(\eta_0, \lambda) = 0.$$

Its solution is obtainable by reasoning similar to that used above. The general solution of the equation (9.9a) is of the form

$$(9.10) \quad v^{(h)}(x, \lambda) = c^{(h)}(\lambda) \mathfrak{Y}^{-1}(x, \lambda) + \int_{\eta_0}^x f(x_1, \lambda) \mathfrak{Y}(x_1, \lambda) \mathfrak{Y}^{-1}(x, \lambda) dx_1,$$

and with this the conditions (9.9b) take the form

$$(9.11) \quad -c^{(h)}(\lambda) - \int_{\eta_0}^{\eta_h} f(x_1, \lambda) \mathfrak{Y}(x_1, \lambda) dx_1 = a_1(\lambda) \mathfrak{B}^{(h)}(\lambda) \mathfrak{Y}(\eta_h, \lambda),$$

$$\sum_{\mu=1}^m c^{(\mu)}(\lambda) = 0, \quad h = 1, 2, \dots, m.$$

An addition of these leads to the evaluation of the parametric vector

$$(9.12) \quad a_1(\lambda) = - \sum_{\mu=1}^m \int_{\eta_0}^{\eta_\mu} f(x_1, \lambda) \mathfrak{Y}(x_1, \lambda) dx_1 \mathfrak{D}^{-1}(\lambda),$$

and in terms of this the vectors $c^{(h)}(\lambda)$ are given by (9.11). The solution, by (9.10), is then explicitly

$$(9.13) \quad v^{(h)}(x, \lambda) = \int_{\eta_h}^x f(x_1, \lambda) \mathfrak{Y}(x_1, \lambda) \mathfrak{Y}^{-1}(x, \lambda) dx_1$$

$$+ \sum_{\mu=1}^m \int_{\eta_0}^{\eta_\mu} f(x_1, \lambda) \mathfrak{G}^{(\mu)}(x, x_1, \lambda) dx_1, \quad h = 1, 2, \dots, m.$$

The problem (9.9), like the problem (9.1), is thus solvable, irrespective of $f(x, \lambda)$, for all values of λ other than the characteristic values.

10. The Green's matrix for a linear x domain. In the classical case of a real variable x , the region of the variable, being a segment of the axis, is not a

two-dimensional domain but a one-dimensional one. It is of some interest on this account to consider somewhat further the more immediate generalization of this case to that in which the domain of x consists of a set of curves in the complex plane which respectively join a point η_0 to the points $\eta_1, \eta_2, \dots, \eta_m$. In the absence of any requirement that these curves be distinct, the configuration is seen to be immediately specializable to the case of a real x and boundary conditions which apply at more than two points of a given interval. The adaptation of the discussion already made to this case of a general curvilinear x domain calls for no modification of the deductions of §7. It permits, however, of an interesting reformulation of the matter of §8, and of an extension of the considerations of §9.

Let the boundary problem adjoint to (7.1) be defined in this case by the equations

$$(10.1a) \quad \mathfrak{z}'(x, \lambda) = -\mathfrak{z}(x, \lambda) \{ \lambda \mathfrak{R}(x) + \mathfrak{Q}(x, \lambda) \},$$

$$\mathfrak{z}(\eta_h, \lambda) + \alpha(\lambda) \mathfrak{B}^{(h)}(\lambda) = 0, \quad h = 1, 2, \dots, m,$$

$$(10.1b) \quad \sum_{\mu=1}^m \mathfrak{z}(\eta_0 + 0\eta_\mu, \lambda) = 0,$$

the solution to exist for a suitable parametric vector $\alpha(\lambda)$, and to consist of a vector $\mathfrak{z}(x, \lambda)$ which satisfies the conditions (10.1b) and solves the equation (10.1a) along each one of the curves constituting the x domain. The symbol $\mathfrak{z}(\eta_0 + 0\eta_h, \lambda)$ is to be interpreted as designating the limit of $\mathfrak{z}(x, \lambda)$ as $x \rightarrow \eta_0$, the approach being along the x curve from η_h . The solution vector $\mathfrak{z}(x, \lambda)$ will in general be discontinuous at the point η_0 , that is, the vectors $\mathfrak{z}(x_0 + 0\eta_h, \lambda)$ will not in general coincide for all h . The deductions of §8 are adapted to this formulation of the adjoint problem without difficulty, being made, in fact, by merely identifying the vector $\mathfrak{z}^{(h)}(x, \lambda)$ of the solution of (8.1), as the solution $\mathfrak{z}(x, \lambda)$ of the problem (10.1) when the variable is on the respective curve from η_0 to η_h .

If x and x_1 are regarded as independent variables, both confined to the given set of curves, a matrix $\mathfrak{G}_1(x, x_1, \lambda)$ is defined by the formula

$$(10.2) \quad \mathfrak{G}_1(x, x_1, \lambda) = \pm \frac{1}{2} \mathfrak{Y}(x, \lambda) \mathfrak{Y}^{-1}(x_1, \lambda),$$

if it is agreed that the plus sign is to apply when x_1 lies on the curve segment which is terminated by the points η_0 and x , while the minus sign is to apply otherwise. The formula

$$(10.3) \quad \mathfrak{G}(x, x_1, \lambda) \equiv \mathfrak{G}_1(x, x_1, \lambda) - \mathfrak{Y}(x, \lambda) \mathfrak{D}^{-1}(\lambda) \sum_{\mu=1}^m \mathfrak{B}^{(\mu)}(\lambda) \mathfrak{G}_1(\eta_\mu, x_1, \lambda),$$

then defines its left-hand member, which will be designated briefly as *the*

Green's matrix. This matrix $\mathfrak{G}(x, x_1, \lambda)$ is related in several ways to the Green's matrices previously defined by the formulas (9.6). Thus, in particular, it will be seen that

$$(10.4) \quad \mathfrak{G}(\eta_\mu, x_1, \lambda) = -\mathfrak{G}^{(h)}(\eta_\mu, x_1, \lambda) + \delta_{\mu,h} \mathfrak{Y}(\eta_\mu, \lambda) \mathfrak{Y}^{-1}(x_1, \lambda),$$

$$\mu, h = 1, 2, \dots, m,$$

whenever x_1 lies on the curve from η_0 to η_h , whereas when x lies on that curve, then

$$(10.5) \quad \mathfrak{G}(x, \eta_0 + 0\eta_\mu, \lambda) = -\mathfrak{G}^{(u)}(x, \eta_0, \lambda) + \delta_{\mu,u} \mathfrak{Y}(x, \lambda) \mathfrak{Y}^{-1}(\eta_0, \lambda).$$

The matrix $\mathfrak{G}(x, x_1, \lambda)$ evidently depends upon x solely by virtue of the occurrence of the matrix $\mathfrak{Y}(x, \lambda)$ as a left-hand factor in the formula (10.3). It follows from this that as a function of x (that is, when x_1 is regarded as fixed) this matrix satisfies the equation (7.1a) along each of the arcs into which the domain of x is divided by the points η_0 and x_1 . Beyond that it is clear from the formula (10.4) and the relation (9.7b) that

$$(10.6) \quad \sum_{\mu=1}^m \mathfrak{B}^{(u)}(\lambda) \mathfrak{G}(\eta_\mu, x_1, \lambda) \equiv \mathfrak{D},$$

namely, that as a function of x it satisfies the condition (7.1b). Formally, therefore, the Green's matrix as a function of its first argument solves the boundary problem (7.1). It fails of being a true solution of that problem because of a discontinuity inherent in it at the point $x = x_1$, for, as is easily verified,

$$(10.7) \quad \mathfrak{G}(x_1 + 0\eta_h, x_1, \lambda) - \mathfrak{G}(x_1 - 0\eta_0, x_1, \lambda) \equiv \mathfrak{F},$$

for any x_1 on the curve from η_0 to η_h .

In an entirely similar manner it will be observed that $\mathfrak{G}(x, x_1, \lambda)$ depends upon x_1 solely by virtue of the presence of the matrix $\mathfrak{Y}^{-1}(x_1, \lambda)$ as a right-hand factor in the formula (10.3). Hence as a function of x_1 (that is, with x fixed) it formally solves the equation (8.1a) along each of the arcs into which the domain of the variable is divided by the points η_0 and x . Since from (10.3) together with (10.5) and (9.7a)

$$(10.8) \quad \mathfrak{G}(x, \eta_h, \lambda) + \mathfrak{Y}(x, \lambda) \mathfrak{D}^{-1}(\lambda) \mathfrak{B}^{(h)}(\lambda) = \mathfrak{D}, \quad h = 1, 2, \dots, m,$$

$$\sum_{\mu=1}^m \mathfrak{G}(x, \eta_0 + 0\eta_\mu, \lambda) = \mathfrak{D},$$

it is seen that as a function of its second argument the Green's matrix is formally a solution of the boundary problem (10.1), with the matrix $\mathfrak{Y}(x, \lambda) \mathfrak{D}^{-1}(\lambda)$ in the role of the parametric matrix. In this instance, as

before, however, it fails to be a true solution because of its discontinuity.

The nonhomogeneous boundary problem (9.1) and also the problem

$$\begin{aligned} v'(x, \lambda) &= -v(x, \lambda) \{ \lambda \mathfrak{R}(x) + \mathfrak{Q}(x, \lambda) \} + f(x, \lambda), \\ v(\eta_h, \lambda) + a_1(\lambda) \mathfrak{B}^{(h)}(\lambda) &= 0, \quad h = 1, 2, \dots, m, \\ \sum_{\mu=1}^m v(\eta_0 + 0\eta_\mu, \lambda) &= 0, \end{aligned} \quad (10.9)$$

which is the reformulation of (9.9), may now be considered, with the $f(x, \lambda)$ as any vectors which are defined merely over the curves of the x domain, and are integrable over these curves. It is easily verified that the solutions of these problems are then given respectively, by the formulas

$$\begin{aligned} u(x, \lambda) &= \sum_{\mu=1}^m \int_{\eta_0}^{\eta_\mu} \mathfrak{G}(x, x_1, \lambda) f(x_1, \lambda) dx_1, \\ v(x, \lambda) &= - \sum_{\mu=1}^m \int_{\eta_0}^{\eta_\mu} f(x_1, \lambda) \mathfrak{G}(x_1, x, \lambda) dx_1. \end{aligned} \quad (10.10)$$

This result is, of course, entirely familiar in its specialization to the case of a real variable with boundary conditions applying at just two points. It seems, however, to be more explicit and compact than any that has heretofore been given even for the case of a real variable, when the boundary conditions are taken to apply at intermediate points as well as at the end points of the interval.

11. On the characteristic values when the boundary conditions apply in a fundamental region. Returning to the discussion as it was left in §9, it will be observed that the deductions of that and the two preceding sections were in the main qualitative, or of a formal nature only. The derivation of more quantitative results requires as a basis some more specific assumptions than those which have heretofore been made. A consideration of the distribution of the characteristic values in the remote part of the λ region, which is now to be undertaken, is, therefore, to be based upon the following addition to the hypotheses of the discussion.

HYPOTHESIS (ii). (a) *The points $\eta_1, \eta_2, \dots, \eta_m$, at which the boundary conditions apply, lie in some fundamental x region, while* (b) *in the part $|\lambda| > N$ of the relative suitable λ region, the matrices $\mathfrak{B}^{(h)}(\lambda)$, ($h=1, 2, \dots, m$), which define the boundary conditions, are analytic and admit of either actual or asymptotic representations of the form*

$$\mathfrak{B}^{(h)}(\lambda) \sim \lambda^\sigma \sum_{k=0}^{\infty} \lambda^{-k} \mathfrak{B}^{(h,k)}, \quad (11.1)$$

in which the matrices $\mathfrak{B}^{(h,k)}$ are constant, and σ is an integer (positive, negative, or zero) such that $\mathfrak{B}^{(h,0)} \neq 0$ for some index h .

By the definition of a fundamental x region and the deductions of §6, the related suitable λ region may be covered by a finite number of subregions Λ , such that while λ remains in any one such, some solution of the matrix equation (2.1) maintains the form (6.10) for all x concerned. With the use of this solution $\mathfrak{Y}(x, \lambda)$, the formula (7.4) yields for the matrix $\mathfrak{D}(\lambda)$ the form

$$(11.2) \quad \mathfrak{D}(\lambda) = \left(\sum_{\mu=1}^m \sum_{r=1}^n w_{i,r}^{(\mu)}(\lambda) p_{r,j}(\eta_{\mu}, \lambda) e^{\lambda R_j(\eta_{\mu})} \right).$$

The determinant $D(\lambda)$, when expanded, is, therefore, given by a formula

$$(11.3) \quad D(\lambda) = \sum_{\alpha} A_{\alpha}(\lambda) e^{\lambda \Omega_{\alpha}},$$

in which

- (a) the index α covers some finite range;
- (b) the symbols Ω_{α} stand for distinct complex constants, which are all included in the set of values which may be obtained from the formula $\sum_{r=1}^n R_r(\eta_{\mu})$ by giving to each index μ , independently one of the values $1, 2, \dots, m$;
- (c) for the coefficient functions,

$$(11.4) \quad A_{\alpha}(\lambda) \neq 0, \quad \text{for each } \alpha.$$

The representability of the coefficient functions $A_{\alpha}(\lambda)$ in a form

$$(11.5) \quad A_{\alpha}(\lambda) \sim \lambda^{\rho_{\alpha}} \sum_{k=0}^{\infty} A_{\alpha,k} \lambda^{-k},$$

follows from (11.1) and (6.11). In this the coefficients $A_{\alpha,k}$ are constants, and the exponents ρ_{α} are integers such that $A_{\alpha,0} \neq 0$, for each α .

The evaluation (11.3) thus obtained depends upon a choice of the solution $\mathfrak{Y}(x, \lambda)$ of the equation (2.1), and the result is valid for a subregion Λ since the form of the solution was specific to such a subregion. As has been previously observed, however, in §7, the determinants $D(\lambda)$ formed from different solutions $\mathfrak{Y}(x, \lambda)$ differ among each other only by factors which are non-vanishing. It may readily be inferred from this that the forms of any specific $D(\lambda)$, formed from a specific solution $\mathfrak{Y}(x, \lambda)$, in different subregions Λ differ from (11.3) at most by such factors. Thus the characteristic values in the entire originally given λ region are simply the zeros of (11.3), that is, the roots of the equation

$$(11.6) \quad \sum_{\alpha} A_{\alpha}(\lambda) e^{\lambda \Omega_{\alpha}} = 0.$$

The left-hand member of this equation may, depending upon the various elements involved in the boundary problem, consist of no terms at all, of a single term, or of more terms than one. The third of these possibilities is that of the greatest interest; the first two are readily disposed of. If the sum consists of no terms at all, the equation (11.6) is vacuous, and imposes no restriction at all upon λ . The boundary problem is accordingly compatible for all λ of the given region. From the formula (11.2) it may be observed that this case inevitably maintains whenever the rank of the matrix

$$\|\mathfrak{B}^{(1)}(\lambda), \mathfrak{B}^{(2)}(\lambda), \dots, \mathfrak{B}^{(m)}(\lambda)\|$$

is less than n . Phrased relatively to the scalar differential system (1), this is merely the statement that the boundary problem is compatible for all λ if the independent boundary conditions are less than n in number. If the number of terms in (11.6) is just one, there are no characteristic values in the domain $|\lambda| > N$. The boundary problem is incompatible for all such λ .

If the left-hand member of (11.6) consists of two or more terms, it is functionally of a structure which is known as an *exponential sum*. The zeros of such a sum are discrete. Their distribution in the λ plane is known,* and may be briefly described as follows. In the complex plane let the points $\bar{\Omega}_\alpha$ (the complex conjugates of the Ω_α) be plotted, and let \bar{P} designate the smallest convex polygon which contains them all in its interior or upon its perimeter. The characteristic values in the λ region in question are all located within a finite number of strips of that region, each strip being bounded by two curves which have asymptotes that are parallel to each other and normal to a side of the polygon \bar{P} . With each of these strips there is associated a pair of constants γ and δ , such that for any choices of $|\lambda_0|$ and Δ , the number of characteristic values which lie in the strip and between the arcs $|\lambda| = |\lambda_0|$, and $|\lambda| = |\lambda_0| + \Delta$, is between $\gamma\Delta - \delta$ and $\gamma\Delta + \delta$.

PART III. THE REPRESENTATION OF ARBITRARY VECTORS

12. Further hypotheses; contours in the λ plane. The considerations which have been set forth in the preceding sections have been based, insofar as they have depended upon the parameter, upon an assumption of the existence merely of some suitable λ region. The results have bearing, therefore, only relative to such regions, even though in specific instances these may constitute but minor portions of the entire λ plane. This does not suffice for the considerations with which the discussion is to continue. For these it is essential, rather, that some qualitative facts be available for all values of λ ,

* Cf. R. E. Langer, *On the zeros of exponential sums and integrals*, Bulletin of the American Mathematical Society, vol. 37 (1931), p. 213.

and that quantitative results be generally applicable to the entire remote portion of the plane, that is, for $|\lambda| > N$. To insure this, the basis of hypotheses must be enlarged, and this is to be done by addition of the following:

HYPOTHESIS (iii). (a) *The differential matrix equation (7.1a) is one for which the entire λ plane is a suitable region; (b) the points $\eta_1, \eta_2, \dots, \eta_m$, at which the boundary conditions apply, lie in an x region which is fundamental relative to the entire λ plane; (c) the elements of the matrices $\mathfrak{B}^{(h)}(\lambda)$, ($h=1, 2, \dots, m$), of (7.1b) are rational functions of λ ; (d) the boundary conditions (7.1b) are such that the expression (11.3) for the determinant $D(\lambda)$ consists of at least two terms.*

Several observations are in order with respect to this hypothesis. To begin with, it will be noted that by virtue of part (a) the further discussion will be restricted to boundary problems of the type (7.1) in which the coefficient matrix \mathfrak{Q} of the differential equation does not involve the parameter λ . This follows from the fact that as a function of λ this matrix has been restricted to be both analytic and bounded over the entire complex plane. In connection with part (c) of the hypothesis, it will be noted that under it the matrices $\mathfrak{B}^{(h)}(\lambda)$ may without any loss of generality be taken to be polynomials in λ . This inference follows from the fact observed in §7 that the matrices $\mathfrak{B}^{(h)}(\lambda)$ may at will be replaced by $\mathfrak{C}_1(\lambda)\mathfrak{B}^{(h)}(\lambda)$ without thereby affecting the content of the boundary problem. The matrix $\mathfrak{C}_1(\lambda)$ can, however, be chosen so that the elements are polynomials in λ , and such as to remove all poles which the elements of the matrices $\mathfrak{B}^{(h)}(\lambda)$ may have in points of the finite λ plane, that is, such that the elements of the matrices $\mathfrak{C}_1(\lambda)\mathfrak{B}^{(h)}(\lambda)$ are integral rational functions of λ . It may be assumed in virtue of this, and it will henceforth be assumed, that such a formal adjustment has been made, and that, therefore, the formulas (11.1) are hereinafter superseded by

$$(12.1) \quad \mathfrak{B}^{(h)}(\lambda) = \lambda^\sigma \sum_{k=0}^{\sigma} \lambda^{-k} \mathfrak{B}^{(h,k)}, \quad h = 1, 2, \dots, m,$$

the matrices $\mathfrak{B}^{(h,k)}$ being still constant, and σ being now a nonnegative integer such that $\mathfrak{B}^{(h,0)} \neq \mathfrak{Q}$ for at least one index value h .

Under Hypothesis (iii) the matrix $\mathfrak{D}(\lambda)$ is analytic over the entire λ plane. The determinant $D(\lambda)$ is, therefore, likewise analytic; hence its zeros, that is, the characteristic values, in any bounded region of the λ plane are finite in number. This applies in particular to the region $|\lambda| \leq N$, whatever the constant N may be. Since for an appropriately large value of N the distribution of the characteristic values in the domain $|\lambda| > N$ is, by virtue of part (d) of the hypothesis, such as is obtained by applying the results of §11 to the whole λ plane, it is seen, in particular, that these values have no finite limit point. They are, therefore, enumerable, and may, in particular, be so enu-

merated that $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3| \leq \dots$. It will be assumed in the following that such an enumeration has been made and will be retained.

Because the characteristic values in the region $|\lambda| > N$ lie in a finite number of strips of the plane, and their densities in these strips are bounded, as was remarked at the end of §11, it is possible to draw in the λ plane certain closed contours which encircle the origin, pass through no characteristic value, and coincide with circles on which $|\lambda|$ is constant, except possibly where they traverse the strips containing the characteristic values. There exists, moreover, an unending sequence of such contours, of which each encloses its predecessor in the sequence, and such that no one of the sequence passes within less than some specifiable positive distance of any characteristic value. Every such contour encloses, of course, only a finite number of characteristic values. If the contours are designated by Γ_κ , with the index κ so assigned as to denote the number of characteristic values enclosed, the following can be shown. There exists a sequence of simple closed contours Γ_κ , as partially described above, for which

(a) the sequence of index values κ is an unbounded increasing sequence of positive integers;

(b) the ratio of κ to the shortest distance from the origin to the contour Γ_κ is bounded;

(c) the ratio of the length of the contour Γ_κ to κ is bounded.

If ρ is used to designate the smallest of the integers ρ_α which occurs in the formulas (11.5) and for which $\bar{\Omega}_\alpha$ is one of the vertices of the polygon \bar{P} described in §11, the function

$$(12.2) \quad \lambda^{-\rho} D(\lambda) e^{-\lambda \bar{\Omega}_\alpha}$$

is an exponential sum whose coefficients are each asymptotic to some nonnegative power of λ . The zeros of this sum, moreover, are simply the characteristic values. Now it is known of such sums* that they remain uniformly bounded from zero for all values of the variable which are uniformly bounded from the roots of the sum. Since λ is so bounded from the characteristic values when it is restricted to vary over the contours of the set Γ_κ as described above, it must be inferred that for λ on such a set of contours the reciprocals of all the functions (12.2) are bounded.

13. The generalized relation of biorthogonality. As has been variously remarked above, and particularly in §7, the matrix $\mathfrak{D}(\lambda)$ is not uniquely determined by the boundary problem, it being in fact a mere matter of adjustment to replace any specific $\mathfrak{D}(\lambda)$ by the matrix (7.6) with any pre-

* Cf. R. E. Langer, *The asymptotic location of the roots of a certain transcendental equation*, these Transactions, vol. 31 (1929), p. 837.

scribed nonsingular analytic matrices \mathfrak{C}_1 and \mathfrak{C}_2 . Deductions which are intrinsic to the boundary problem are, of course, independent of such adjustment. Their derivation may, however, be simpler with a fortunate adjustment than with a contrary one, and this is the case in the discussions of the present and the following sections. There are, in other words, advantages of simplicity to be gained by a suitable normalization of the matrix $\mathfrak{D}(\lambda)$.

Let λ_β be any characteristic value, and for generality let its multiplicity be denoted by s . Whatever the adjustment of the problem, the matrix $\mathfrak{D}(\lambda)$, being analytic, has elements which are expansible in power series in $(\lambda - \lambda_\beta)$. The initial segments of these series, extending to the terms in $(\lambda - \lambda_\beta)^s$, are polynomials in $(\lambda - \lambda_\beta)$ of degree s . If their matrix is designated by $\mathfrak{P}(\lambda)$, it is clear that

$$(13.1) \quad \mathfrak{D}(\lambda) \equiv \mathfrak{P}(\lambda) + (\lambda - \lambda_\beta)^{s+1} \mathfrak{D}^{(1)}(\lambda),$$

with $\mathfrak{D}^{(1)}(\lambda)$ analytic at the value λ_β . Now since \mathfrak{P} is a polynomial matrix, it will, under a suitable adjustment in the sense above, appear in its canonical form

$$(13.2) \quad \mathfrak{P}(\lambda) \equiv (\delta_{i,j} p_i(\lambda)),$$

in which each element p_i is a polynomial in $(\lambda - \lambda_\beta)$, with unity as the coefficient of the lowest power of $(\lambda - \lambda_\beta)$ which is actually present, and each element p_i a factor of its successor p_{i+1} .* The adjustment of the problem for which (13.1) and (13.2) obtain will be assumed throughout the immediately following discussion. The matrix $\mathfrak{Y}(x, \lambda)$ is thereby in part determined.

For generality let it be assumed now that the index of the characteristic value λ_β is r . Then $(n-r)$ is the rank of the matrix $\mathfrak{D}(\lambda_\beta)$, and since it is clear from (13.1) that $\mathfrak{P}(\lambda_\beta)$ is of the same rank, it follows that

$$(13.3) \quad p_i(\lambda_\beta) = \begin{cases} 1, & \text{for } i \leq n-r, \\ 0, & \text{for } i > n-r. \end{cases}$$

Since the zero of $D(\lambda)$ at λ_β is of the same multiplicity as the zero of the determinant of $\mathfrak{P}(\lambda)$, that is, of $\prod_{i=1}^n p_i(\lambda)$, whereas each of the factors $p_i(\lambda)$, ($i = n-r+1, \dots, n$), has a zero at λ_β , it is clear that the multiplicity s of this value is at least as great as its index r , a fact which was stated in §7. Now due to the rank of $\mathfrak{D}(\lambda_\beta)$ there are precisely r linearly independent vectors c which satisfy the equation (7.3) at λ_β , and each of these leads through (7.2) to a characteristic solution of the boundary problem. However, since all elements in the last r columns of $\mathfrak{D}(\lambda_\beta)$ are zero, it is seen at once that each of the

* Cf. M. Bôcher, *Introduction to Higher Algebra*.

vectors which, with j fixed at one of the values $n-r+1, \dots, n$, has the components $\delta_{i,j}$, does serve as a solution c of the equation (7.3). The formula (7.2), therefore, gives as characteristic solutions

$$(13.4) \quad \eta^{(k,\beta)}(x) \equiv \mathcal{Y}(x, \lambda_\beta)(\delta_{i,n-r+k}), \quad k = 1, 2, \dots, r,$$

and these solutions are thus seen to be given precisely by the last r columns of the matrix $\mathcal{Y}(x, \lambda_\beta)$.

Again, at λ_β there are precisely r linearly independent vectors a which solve the equation (8.4), and due to the fact that all elements in the last r rows of the matrix $\mathcal{D}(\lambda_\beta)$ are zeros, it is clear that with i fixed at any one of the values $n-r+1, \dots, n$, the vector with components $\delta_{i,j}$ is such a one. Through the formulas (8.3) and (8.2), it follows then that the formulas

$$(13.5) \quad \mathfrak{z}^{(k,\beta,h)}(x) \equiv -(\delta_{n-r+k,i})\mathfrak{B}^{(h)}(\lambda_\beta)\mathcal{Y}(\eta_h, \lambda_\beta)\mathcal{Y}^{-1}(x, \lambda_\beta), \quad h = 1, 2, \dots, m,$$

yield, for each k on the range $1, 2, \dots, r$, a characteristic solution of the problem (8.1). These solutions are thus given by the last r rows of the matrices

$$(13.6) \quad -\mathfrak{B}^{(h)}(\lambda_\beta)\mathcal{Y}(\eta_h, \lambda_\beta)\mathcal{Y}^{-1}(x, \lambda_\beta), \quad h = 1, 2, \dots, m.$$

Let λ be regarded now as distinct from λ_β , and let $\mathfrak{z}^{(k,\beta,h)}(x)$ be any one of the characteristic solutions (13.5). The obvious relation

$$\begin{aligned} \sum_{\mu=1}^m \int_{\eta_0}^{\eta_\mu} \{ \mathfrak{z}^{(k,\beta,h)}(x)\mathcal{Y}'(x, \lambda) + \mathfrak{z}^{(k,\beta,h)'}(x)\mathcal{Y}(x, \lambda) \} dx \\ = \sum_{\mu=1}^m \{ \mathfrak{z}^{(k,\beta,h)}(\eta_\mu)\mathcal{Y}(\eta_\mu, \lambda) - \mathfrak{z}^{(k,\beta,h)}(\eta_0)\mathcal{Y}(\eta_0, \lambda) \}, \end{aligned}$$

assumes, then, because of the relations (8.1) and (7.1a), the form

$$(13.7) \quad (\lambda - \lambda_\beta) \sum_{\mu=1}^m \int_{\eta_0}^{\eta_\mu} \mathfrak{z}^{(k,\beta,h)}(x_1)\mathfrak{R}(x_1)\mathcal{Y}(x_1, \lambda)dx_1 = -\mathfrak{a}^{(k,\beta)} \sum_{\mu=1}^m \mathfrak{B}^{(\mu)}(\lambda_\beta)\mathcal{Y}(\eta_\mu, \lambda),$$

$$\mathfrak{a}^{(k,\beta)} = (\delta_{n-r+k,i}).$$

Now since the matrices $\mathfrak{B}^{(\mu)}(\lambda)$ are polynomials of degree σ , as shown by (12.1), it is clear that with an arbitrary choice of $(\tau+1)$ as a nonnegative integer, the left-hand member of the formula

$$(13.8) \quad \left(\frac{\lambda}{\lambda_\beta} \right)^{\tau+1} \{ \mathfrak{B}^{(\mu)}(\lambda) - \mathfrak{B}^{(\mu)}(\lambda_\beta) \} = (\lambda - \lambda_\beta) \sum_{h=r+1}^{\tau+\sigma} \lambda^h \mathfrak{B}^{(\mu,h)}(\lambda_\beta)$$

is a polynomial in λ which vanishes at λ_β . Its structure is, therefore, such as is shown on the right of (13.8), and this relation may be looked upon as defining the matrices

$$\mathfrak{B}^{(\mu, h)}(\lambda), \quad h = \tau + 1, \tau + 2, \dots, \tau + \sigma.$$

It is likewise seen that

$$(13.9) \quad \left\{ \left(\frac{\lambda}{\lambda_\beta} \right)^{\tau+1} - 1 \right\} \mathfrak{B}^{(\mu)}(\lambda_\beta) = (\lambda - \lambda_\beta) \sum_{h=0}^{\tau} \lambda^h \mathfrak{B}^{(\mu, h)}(\lambda_\beta),$$

with

$$(13.10) \quad \mathfrak{B}^{(\mu, h)}(\lambda) \equiv \lambda^{-h-1} \mathfrak{B}^{(\mu)}(\lambda), \quad h = 0, 1, \dots, \tau.$$

If the relations (13.8) and (13.9) are added, and the sum is multiplied on the right by $\mathfrak{Y}(\eta_\mu, \lambda)$, it is found, on recalling (7.4) that

$$\left(\frac{\lambda}{\lambda_\beta} \right)^{\tau+1} \mathfrak{D}(\lambda) - \sum_{\mu=1}^m \mathfrak{B}^{(\mu)}(\lambda_\beta) \mathfrak{Y}(\eta_\mu, \lambda) = (\lambda - \lambda_\beta) \sum_{\mu=1}^m \sum_{h=0}^{\tau+\sigma} \mathfrak{B}^{(\mu, h)}(\lambda_\beta) \lambda^h \mathfrak{Y}(\eta_\mu, \lambda).$$

In virtue of this the relation (13.7) may be written

$$(13.11) \quad \sum_{\mu=1}^m \left\{ - \int_{\eta_0}^{\eta_\mu} \delta^{(k, \beta, \mu)}(x_1) \mathfrak{R}(x_1) \mathfrak{Y}(x_1, \lambda) dx_1 + \sum_{h=0}^{\tau+\sigma} a^{(k, \beta)} \mathfrak{B}^{(\mu, h)}(\lambda_\beta) \lambda^h \mathfrak{Y}(\eta_\mu, \lambda) \right\} \\ = \left(\frac{\lambda}{\lambda_\beta} \right)^{\tau+1} \frac{1}{\lambda - \lambda_\beta} a^{(k, \beta)} \mathfrak{D}(\lambda).$$

Let λ be taken now as any characteristic value, say λ_γ , distinct from λ_β , and let $\eta^{(q, \gamma)}(x)$ be any corresponding characteristic solution of the problem (7.1). There exists then a vector c which satisfies the equation (7.3), and for which the left-hand member of (7.2) is $\eta^{(q, \gamma)}(x)$. If the relation (13.11) is multiplied by this vector c upon the right it is found as a result that

$$(13.12) \quad \sum_{\mu=1}^m \left\{ - \int_{\eta_0}^{\eta_\mu} \delta^{(k, \beta, \mu)}(x_1) \mathfrak{R}(x_1) \eta^{(q, \gamma)}(x_1) dx_1 \right. \\ \left. + \sum_{h=0}^{\tau+\sigma} a^{(k, \beta)} \mathfrak{B}^{(\mu, h)}(\lambda_\beta) \lambda_\gamma^h \eta^{(q, \gamma)}(\eta_\mu) \right\} = 0.$$

The vector $a^{(k, \beta)} \mathfrak{D}(\lambda)$ which occurs on the right of (13.11) is represented simply by the $(n-r+k)$ th row in the matrix $\mathfrak{D}(\lambda)$. Every element of this row has a zero at λ_β . The right-hand member of (13.11) is, therefore, analytic at λ_β if properly defined there. Consider now the case in which the index r of the value λ_β equals its multiplicity s . Since the zero of the product of r factors $\prod_{i=n-r+1}^n p_i(\lambda)$ is precisely of the multiplicity r , each factor has a zero of precisely the first order, that is, each of the elements $p_i(\lambda)$ of (13.2) for which $i > n-r$ is a polynomial in $(\lambda - \lambda_\beta)$ of which the term of lowest degree is precisely $(\lambda - \lambda_\beta)$. It is clear from this, in virtue of (13.1), that

$$(13.13) \quad \lim_{\lambda \rightarrow \lambda_\beta} \frac{1}{\lambda - \lambda_\beta} \alpha^{(k, \beta)} \mathfrak{D}(\lambda) = \alpha^{(k, \beta)}.$$

Since

$$\alpha^{(k, \beta)}(\delta_{i, n-r+q}) = \delta_{k, q},$$

it is seen that if the relation (13.11) is multiplied on the right by the vector $(\delta_{i, n-r+q})$ with $q = 1, 2, \dots, r$, and λ is allowed to approach λ_β , the limiting form is, in virtue of (13.4),

$$(13.14) \quad \sum_{\mu=1}^m \left\{ - \int_{\eta_0}^{\eta_\mu} \delta^{(k, \beta, \mu)}(x_1) \Re(x_1) \eta^{(q, \beta)}(x_1) dx_1 + \sum_{h=0}^{\tau+q} \alpha^{(k, \beta)} \mathfrak{B}^{(\mu, h)}(\lambda_\beta) \lambda_\beta^h \eta^{(q, \beta)}(\eta_\mu) \right\} = \delta_{k, q}.$$

This deduction does not follow if the multiplicity s of the value λ_β exceeds its index r . For, in that case at least one of the elements $p_i(\lambda)$, with $i > n-r$, has at λ_β a zero of order higher than the first. For at least one value of k , therefore, the left-hand member of the relation (13.13) is zero. The relation (13.14) is, therefore, invalid, its left-hand member being zero irrespective of q when k has certain values.

The results of this section, as involved in the formulas (13.12) and (13.14) may be formulated as follows: If λ_β is any characteristic value whose index equals its multiplicity, then

$$(13.15) \quad \sum_{\mu=1}^m \left\{ - \int_{\eta_0}^{\eta_\mu} \delta^{(k, \beta, \mu)}(x_1) \Re(x_1) \eta^{(q, \gamma)}(x_1) dx_1 + \sum_{h=0}^{\tau+q} (\delta_{n-r+k, i}) \mathfrak{B}^{(\mu, h)}(\lambda_\beta) \lambda_\beta^h \eta^{(q, \gamma)}(\eta_\mu) \right\} = \delta_{\beta, \gamma} \cdot \delta_{k, q},$$

where r is the index of λ_β , and τ is any integer not less than -1 . If the multiplicity of λ_β exceeds its index, the relation (13.15) fails for at least one value of k , the right-hand member of the relation for such k being 0 for all γ and q .

This result must evidently be looked upon as the generalization of the relation of ordinary or weighted biorthogonality which is familiarly a property of the set of characteristic solutions of adjoint boundary problems in the classical specialized cases. Evidently the set may be normalized, in the sense that for $\gamma = \beta$ and $q = k$ the right-hand member of (13.15) is unity, whenever the characteristic values all have indices equal to their multiplicities, whereas complete normalization is impossible when this condition is not fulfilled.

14. **The residues of the Green's matrices.** The matrix $\mathfrak{D}^{-1}(\lambda)$, as has been observed, is analytic over the λ plane except at the characteristic values, where

it has poles. It is appropriate at this point to turn the considerations to the deduction of the residues of the Green's matrices (9.6) at these characteristic values. For this purpose the residue of a matrix, say of $\mathfrak{G}^{(h)}(x, x_1, \lambda)$, at the pole λ_β , will be designated by the symbol $\text{res}_\beta \mathfrak{G}^{(h)}(x, x_1)$. For convenience the choice of the matrix $\mathfrak{Y}(x, \lambda)$ and the adjustment of the boundary problem will be taken, as in the preceding section, to be such that the matrix $\mathfrak{D}(\lambda)$ is in the canonical form (13.1), (13.2). The discussion will be concerned with any characteristic value λ_β whose multiplicity and index are equal.

If the index of λ_β is r , the matrix $\mathfrak{D}(\lambda)$, as has been seen, has $(\lambda - \lambda_\beta)$ as a multiple factor of each element not upon its principal diagonal, while this function is a simple factor of the diagonal elements of the last r columns and is not a factor of the diagonal elements of the first $(n-r)$ columns. The coefficient of the lowest power of $(\lambda - \lambda_\beta)$ occurring in any diagonal element, it will be recalled, is unity. From this it is seen at once that with proper definition at λ_β , and in terms of the matrices (5.9), the matrix

$$(14.1) \quad \mathfrak{G}(\lambda) \equiv \mathfrak{D}(\lambda) \left\{ \sum_{l=1}^{n-r} \mathfrak{Z}_{l,l} + \frac{1}{\lambda - \lambda_\beta} \sum_{l=n-r+1}^n \mathfrak{Z}_{l,l} \right\}$$

is analytic and nonsingular at λ_β . Its elements are polynomials in $(\lambda - \lambda_\beta)$ of which the term of zero degree is precisely $\delta_{i,j}$. The formulas $\mathfrak{G}^{-1}(\lambda_\beta) = \mathfrak{Z}$, and

$$(14.2) \quad \mathfrak{D}^{-1}(\lambda) \equiv \left\{ \sum_{l=1}^{n-r} \mathfrak{Z}_{l,l} + \frac{1}{\lambda - \lambda_\beta} \sum_{l=n-r+1}^n \mathfrak{Z}_{l,l} \right\} \mathfrak{G}^{-1}(\lambda),$$

lead directly to the result

$$(14.3) \quad \text{res}_\beta \mathfrak{D}^{-1} = \sum_{l=n-r+1}^n \mathfrak{Z}_{l,l}.$$

Now from the formula (9.6) and the fact that the poles of $\mathfrak{D}^{-1}(\lambda)$ as shown by (14.2) are of the first order, it follows that

$$(14.4) \quad \text{res}_\beta \mathfrak{G}^{(h)}(x, x_1) = \mathfrak{Y}(x, \lambda_\beta) \{ \text{res}_\beta \mathfrak{D}^{-1} \} \mathfrak{B}^{(h)}(\lambda_\beta) \mathfrak{Y}(\eta_h, \lambda_\beta) \mathfrak{Y}^{-1}(x_1, \lambda_\beta),$$

$$h = 1, 2, \dots, m.$$

However, the formulas (13.4) and (13.5) yield readily the fact that for $l > n-r$

$$\mathfrak{Y}(x, \lambda_\beta) \mathfrak{Z}_{l,l} \mathfrak{B}^{(h)}(\lambda_\beta) \mathfrak{Y}(\eta_h, \lambda_\beta) \mathfrak{Y}^{-1}(x_1, \lambda_\beta) \equiv - (y_i^{(l+r-n, \beta)}(x) z_j^{(l+r-n, \beta, h)}(x_1)),$$

in which the components of the characteristic vectors (13.4) and (13.5) have been designated, respectively, by $y_i^{(k, \beta)}(x)$, ($i=1, 2, \dots, n$) and $c_j^{(k, \beta, h)}(x_1)$, ($j=1, 2, \dots, n$). It follows, on substituting (14.3) into (14.4) that

$$(14.5) \quad \text{res}_\beta \mathfrak{G}^{(h)}(x, x_1) \equiv - \sum_{k=1}^r (y_i^{(k, \beta)}(x) z_j^{(k, \beta, h)}(x_1)), \quad h = 1, 2, \dots, m.$$

The residues of the Green's matrices have thus been explicitly evaluated for all characteristic values whose multiplicities and indices are the same. This, of course, includes in particular all the simple characteristic values.

15. The formal expansion of an arbitrary vector. Let the consideration be turned now to an infinite series

$$(15.1) \quad \sum_{\gamma=1}^{\infty} \sum_{q=1}^{r_{\gamma}} f_{q,\gamma} \eta^{(q,\gamma)}(x),$$

in which r_{γ} is the index of the characteristic value λ_{γ} ; the $\eta^{(q,\gamma)}(x)$ are characteristic solutions; the $f_{q,\gamma}$ are scalar constants; and x varies over some fundamental region of the x plane that contains the points $\eta_0, \eta_1, \dots, \eta_m$, and in which the coefficient matrix $\mathfrak{R}(x)$ in the equation (7.1a) is nonsingular. The existence of such an x region is an assumption. If the coefficients $f_{q,\gamma}$ are such that the series converges uniformly (a tentative heuristic assumption), say to the vector $f(x)$, the term by term differentiation of the series is permissible, and with the use of the equation (7.1a) a process is evident which by repetition leads to the sequence of relations

$$(15.2) \quad \sum_{\gamma=1}^{\infty} \sum_{q=1}^{r_{\gamma}} f_{q,\gamma} \lambda_{\gamma}^h \eta^{(q,\gamma)}(x) = f^{(h)}(x), \quad h = 0, 1, 2, \dots,$$

in which

$$(15.3) \quad f^{(0)}(x) \equiv f(x), \quad f^{(h)}(x) \equiv \mathfrak{R}^{-1}(x) \{ f^{(h-1)'}(x) - \mathfrak{Q}(x) f^{(h-1)}(x) \}, \\ h = 1, 2, \dots.$$

In particular, it follows from this that

$$(15.4) \quad \sum_{\gamma=1}^{\infty} \sum_{q=1}^{r_{\gamma}} f_{q,\gamma} \lambda_{\gamma}^h \eta^{(q,\gamma)}(\eta_{\mu}) = f^{(\mu,h)}, \quad \mu = 1, 2, \dots, m; h = 0, 1, 2, \dots,$$

in which, evidently,

$$(15.5) \quad f^{(\mu,h)} = f^{(h)}(\eta_{\mu}).$$

By the relations (15.2) and (15.4) and the fact that the series involved are integrable term by term, it is seen, then, that

$$\begin{aligned} \sum_{\mu=1}^m \left\{ - \int_{\eta_0}^{\eta_{\mu}} \mathfrak{z}^{(k,\beta,\mu)}(x_1) \mathfrak{R}(x_1) f(x_1) dx_1 + \sum_{h=0}^{\tau+\sigma} (\delta_{n-r+k,j}) \mathfrak{B}^{(\mu,h)}(\lambda_{\beta}) f^{(\mu,h)} \right\} \\ = \sum_{\gamma=1}^{\infty} \sum_{q=1}^{r_{\gamma}} f_{q,\gamma} \sum_{\mu=1}^m \left\{ - \int_{\eta_0}^{\eta_{\mu}} \mathfrak{z}^{(k,\beta,\mu)}(x_1) \mathfrak{R}(x_1) \eta^{(q,\gamma)}(x_1) dx_1 \right. \\ \left. + \sum_{h=0}^{\tau+\sigma} (\delta_{n-r+k,j}) \mathfrak{B}^{(\mu,h)}(\lambda_{\beta}) \lambda_{\gamma}^h \eta^{(q,\gamma)}(\eta_{\mu}) \right\}, \end{aligned}$$

for any choice of k and β . In this relation, however, the series on the right reduces to the single term given by $\gamma = \beta$ and $q = k$, in virtue of (13.15). The excepted term is by (13.15) precisely $f_{k,\beta}$, provided the index and multiplicity of the characteristic value λ_β are equal. If the index is less than the multiplicity, on the other hand, then for at least one value of k this excepted term is zero like the rest. In the former case the result is

$$(15.6) \quad f_{k,\beta} = \sum_{\mu=1}^m \left\{ - \int_{\eta_0}^{\eta_\mu} \delta^{(k,\beta,\mu)}(x_1) \Re(x_1) \bar{f}(x_1) dx_1 \right. \\ \left. + \sum_{h=0}^{r+\sigma} (\delta_{n-r+k,j}) \mathfrak{B}^{(\mu,h)}(\lambda_\beta) \bar{f}^{(\mu,h)} \right\},$$

while in the latter case the scheme leads to no evaluation of $f_{k,\beta}$.

In the case of a boundary problem for which each characteristic value is of index equal to its multiplicity—and in the following deduction this case will be assumed—the series (15.1) is completely explicit in virtue of the formula (15.6). The terms of this series may be expressed as residues, as will now be shown.

If with l designating any nonnegative integer the formula (15.6) is multiplied on the right by $\lambda_\beta^l \eta^{(k,\beta)}(x)$, then in virtue of the relations

$$\delta^{(k,\beta,\mu)}(x_1) \Re(x_1) \bar{f}(x_1) \eta^{(k,\beta)}(x) \equiv (y_i^{(k,\beta)}(x) z_j^{(k,\beta,\mu)}(x_1)) \Re(x_1) \bar{f}(x_1), \\ (\delta_{n-r+k,j}) \mathfrak{B}^{(\mu,h)}(\lambda_\beta) \bar{f}^{(\mu,h)} \eta^{(k,\beta)}(x) \equiv \mathfrak{V}(x, \lambda_\beta) \mathfrak{F}_{n-r+k, n-r+k} \mathfrak{B}^{(\mu,h)}(\lambda_\beta) \bar{f}^{(\mu,h)},$$

the result obtained is

$$f_{k,\beta} \lambda_\beta^l \eta^{(k,\beta)}(x) \equiv \sum_{\mu=1}^m \lambda_\beta^l \left\{ - \int_{\eta_0}^{\eta_\mu} (y_i^{(k,\beta)}(x) z_j^{(k,\beta,\mu)}(x_1)) \Re(x_1) \bar{f}(x_1) dx_1 \right. \\ \left. + \mathfrak{V}(x, \lambda_\beta) \mathfrak{F}_{n-r+k, n-r+k} \sum_{h=0}^{r+\sigma} \mathfrak{B}^{(\mu,h)}(\lambda_\beta) \bar{f}^{(\mu,h)} \right\}.$$

Because of (14.5) and (14.3) this leads to

$$(15.7) \quad \sum_{k=1}^{r\beta} f_{k,\beta} \lambda_\beta^l \eta^{(k,\beta)}(x) \equiv \text{res}_\beta \sum_{\mu=1}^m \lambda^l \left\{ \int_{\eta_0}^{\eta_\mu} \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \Re(x_1) \bar{f}(x_1) dx_1 \right. \\ \left. + \mathfrak{V}(x, \lambda) \mathfrak{D}^{-1}(\lambda) \sum_{h=0}^{r+\sigma} \mathfrak{B}^{(\mu,h)}(\lambda) \bar{f}^{(\mu,h)} \right\}.$$

The series (15.1) may, therefore, be looked upon as an infinite series of residues which are contributed by poles at the characteristic values.

The deductions of this section thus far were based upon certain assumptions, that were signalized as tentative, concerning the coefficients of the

series (15.1), and concerning the characteristic values. They were also based upon the formulas (15.5) and (15.3). Quite independently of the deduction given, however, and with any set of vectors

$$f(x), \quad f^{(\mu, h)}, \quad h = 0, 1, 2, \dots, \tau + \sigma; \mu = 1, 2, \dots, m,$$

of which the first is analytic and the others constant, the right-hand member of (15.7) is specific. This is so, in particular, if the choice

$$\begin{aligned} f^{(\mu, h)} &= f^{(h)}(\eta_\mu), & h &= 0, 1, 2, \dots, \tau, \\ f^{(\mu, h)} &= 0, & h &= \tau + 1, \dots, \tau + \sigma, \end{aligned}$$

is made. In this case the series of right-hand members of (15.7) reduces, because of (13.10), and (9.7b) to

$$\begin{aligned} \mathfrak{f}^{(l)}(x) &\equiv \sum_{\beta=0}^{\infty} \operatorname{res}_{\beta} \sum_{\mu=1}^m \lambda^l \left\{ \int_{\eta_0}^{\eta_\mu} \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \Re(x_1) f(x_1) dx_1 \right. \\ (15.8) \quad &\quad \left. + \mathfrak{G}^{(\mu)}(x, \eta_\mu, \lambda) \sum_{h=0}^{\tau} \lambda^{-h-1} f^{(h)}(\eta_\mu) \right\}, \quad l = 0, 1, 2, \dots \end{aligned}$$

The series of this set, $\mathfrak{f}^{(0)}(x)$, $\mathfrak{f}^{(1)}(x)$, and so on, will be referred to briefly hereinafter as the formal expansions of the vector $f(x)$. It will be observed that to specify them the integer τ must be given, for inasmuch as the chosen set of vectors depends upon τ , the result (15.8) does so likewise. Arrived at in this manner, the questions of convergence of the formal expansions, or of their values in the event of convergence, remain, of course, entirely open. The continuing discussion is designed to bear upon them.

In §12 an ordering of the characteristic values in an order of non-decreasing absolute magnitude was agreed upon, and the existence of the sequence of contours in the λ plane was deduced, the contour Γ_κ of this sequence enclosing the origin and precisely the first κ of the characteristic values. If the symbol res_0 in (15.8) is interpreted to signify the residue at the origin, it is at once clear that with any fixed l the vector

$$\begin{aligned} \mathfrak{f}_\kappa^{(l)}(x) &= \frac{1}{2\pi i} \int_{\Gamma_\kappa} \sum_{\mu=1}^m \left\{ \int_{\eta_0}^{\eta_\mu} \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \Re(x_1) f(x_1) dx_1 \right. \\ (15.9) \quad &\quad \left. + \mathfrak{G}^{(\mu)}(x, \eta_\mu, \lambda) \sum_{h=0}^{\tau} \lambda^{-h-1} f^{(h)}(\eta_\mu) \right\} \lambda^l d\lambda \end{aligned}$$

represents the sum of the first $\kappa+1$ terms of the respective series (15.8). The form of the right-hand member of (15.9) may be somewhat modified, with advantage to the analysis which is to be applied to it. Since the integrand is analytic, the several paths of integration as to x_1 may be chosen at pleasure,

and hence may, in particular, be chosen to pass through the point x , and to coincide from the point η_0 to x . The integrations over this common path contribute to the formula (15.9) the value

$$\frac{1}{2\pi i} \int_{\Gamma_k} \int_{\eta_0}^x \sum_{\mu=1}^m \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \mathfrak{R}_1(x_1) f(x_1) dx_1 \lambda^l d\lambda.$$

This, however, is zero, since by the relation (9.7a) the integrand is seen to be analytic everywhere within the contour Γ_k . The formula (15.9) may, therefore, be written alternatively as

$$(15.10) \quad \mathfrak{G}_k^{(l)}(x) = \frac{1}{2\pi i} \int_{\Gamma_k} \sum_{\mu=1}^m \left\{ - \int_{\eta_0}^x \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \mathfrak{R}_1(x_1) f(x_1) dx_1 \right. \\ \left. + \mathfrak{G}^{(\mu)}(x, \eta_\mu, \lambda) \sum_{h=0}^r \lambda^{-h-1} f^{(h)}(\eta_\mu) \right\} \lambda^l d\lambda, \quad l = 0, 1, 2, \dots.$$

16. Regularity of a boundary problem. Under Hypothesis (iii) of §12, there exists for the differential equation (7.1a) a fundamental x region relative to the entire λ plane, and this region contains the points $\eta_1, \eta_2, \dots, \eta_m$, at which the boundary conditions of the problem (7.1) apply. The variable x has been taken in such a region. Hence the λ plane may be thought of as covered by a finite number of λ sectors, in each of which some solution $\mathfrak{Y}(x, \lambda)$ of the equation (7.1a) maintains the form (6.10), (6.11). When formed from that solution in the respective sector, the matrix $\mathfrak{D}(\lambda)$ has the form (11.2). From this the structure of the determinant $D(\lambda)$, or of any of its minors, may be deduced. The former has already been done in §11, the result (11.3) being valid under present hypotheses in the appropriate λ sector.

Consider then $D_{r,c}(\lambda)$, the cofactor of the element in the r th row and c th column of $D(\lambda)$. From the formula (11.2) this is seen to be, when completely expanded,

$$D_{r,c}(\lambda) = \sum_{\mu_1=1}^m \cdots \sum_{\mu_{c-1}=1}^m \sum_{\mu_{c+1}=1}^m \cdots \sum_{\mu_n=1}^m b_{c,r}(\lambda, \mu_1, \dots, \mu_{c-1}, \mu_{c+1}, \dots, \mu_n) \\ \cdot \exp \left\{ \sum_{\substack{h=1 \\ h \neq c}}^n R_h(\eta_{\mu_h}) \right\},$$

in which $b_{c,r}(\lambda, \mu_1, \dots, \mu_{c-1}, \mu_{c+1}, \dots, \mu_n)$ is the cofactor of the (r, c) th element in the matrix

$$\left(\sum_{j=1}^n w_{i,r}^{(\mu_j)}(\lambda) p_{r,j}(\eta_{\mu_j}, \lambda) \right).$$

in which

$$(16.5) \quad \Psi_{\beta,l,k}^{(\mu)}(x) \equiv \Phi_l^{(\beta)} + R_l(x) + R_k(\eta_\mu),$$

$$(16.6) \quad \mathfrak{S}_{\beta,l,k}^{(\mu)}(x, x_1, \lambda) \equiv \mathfrak{P}(x, \lambda) \mathfrak{Y}_{l,l}^{(\beta)}(\lambda) \mathfrak{B}^{(\mu)}(\lambda) \mathfrak{P}(\eta_\mu, \lambda) \mathfrak{Y}_{k,k}^{-1}(x_1, \lambda).$$

From (16.2), together with the formulas (6.11) and (12.1), it is seen that when $|\lambda| > N$, the matrices (16.6) admit of representations of a form

$$(16.7) \quad \mathfrak{S}_{\beta,l,k}^{(\mu)}(x, x_1, \lambda) \sim \lambda^\theta \sum_{\gamma=1}^{\infty} \lambda^{-\gamma} \mathfrak{S}_{\beta,l,k}^{(\mu,\gamma)}(x, x_1),$$

in which θ is an integer and the matrices on the right are analytic in x and x_1 .

The formula (16.4) was derived on the assumption that λ remains in a sector of the λ plane. However, since the Green's matrices are independent of the choice of the solution $\mathfrak{Y}(x, \lambda)$ from which they are formed, the result is independent of the sector, that is, the formula is valid for all λ .

If x is thought of now as fixed, and x_1 is taken as the variable, it is conceivable that for some choices of μ and k the values given by (16.5) with different indices β, l may not all be distinct. In such case the same exponential occurs in different terms of certain of the sums of (16.4), and a simplification of the respective formulas is achievable by collecting such terms, and omitting from the results any such collections of terms of which the resultant coefficients reduce to the matrix \mathfrak{D} .

Let ζ be taken as a complex variable, and in the plane of ζ let the points Ω_α which are defined by the formula (11.3) be plotted. Then let P designate the smallest convex polygon in the ζ plane, which contains all of these points in its interior or upon its perimeter. For any chosen and fixed value of x , the relations

$$(16.8) \quad \zeta = \Psi_{\beta,l,k}^{(\mu)}(x) - R_k(x_1)$$

define, for each set of indices μ, β, l, k , an analytic map of any configuration in the x_1 plane upon a corresponding configuration in the ζ plane. This latter may or may not in any specific instance fall into the interior of the polygon P , and since x enters into the definition of the transformation in the role of a parameter, this will depend to some extent upon the value x which is in question. With this in mind, the following will be made as a definition.

A boundary problem (7.1) will be defined to be regular as to the point x if (a) the matrix $\mathfrak{R}(x)$ is nonsingular, and if (b) for each set of indices μ, β, l, k , to which there corresponds a term of the (simplified) sums (16.4), there exists in the fundamental region which is the domain of the variable x_1 , some curve joining the point η_μ with the point x which maps under the respective transformation (16.8) into a locus no point of which lies outside of the polygon P .

The condition (a) for regularity is obviously fulfilled at all points of some neighborhood of any point at which it holds, due to the analyticity of the matrix $\mathfrak{R}(x)$. On the other hand, the condition (b) may apparently be fulfilled relative to a point but not relative to neighboring points. This would be so in the case that its fulfillment at x is ascribable to a simplification of the sums in (16.4); for such simplifications are evidently possible only for isolated x values. In suitable cases, however, the condition (b) also may be fulfilled relative to all points of a region. We therefore agree that:

A boundary problem (7.1) is to be designated as regular as to a region of the plane if it is regular as to each point of that region.

17. The convergence of the formal expansions at points of regularity.

If x is taken as a point relative to which the boundary problem is regular, the arbitrarily chosen analytic vector $f(x)$ has associated with it the set of vectors $f^{(h)}(x)$, ($h=1, 2, \dots$), given by the recurrence formula (15.3). Since the Green's matrices as functions of x_1 all satisfy the differential equation (8.1a) it is easily verified that the relations

$$(17.1) \quad \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \mathfrak{R}(x_1) f(x_1) \equiv - \frac{\partial}{\partial x_1} \left\{ \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \sum_{h=0}^{\tau} \lambda^{-h-1} f^{(h)}(x_1) \right\} \\ + \lambda^{-\tau-1} \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \mathfrak{R}(x_1) f^{(\tau+1)}(x_1)$$

are identities, and are valid with any choice of τ as a nonnegative integer. They lead, with the use of the formula (9.7a), to the evaluation

$$(17.2) \quad \sum_{\mu=1}^m \left\{ - \int_{\eta_{\mu}}^x \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \mathfrak{R}(x_1) f(x_1) dx_1 \right. \\ = \sum_{h=0}^{\tau} \lambda^{-h-1} f^{(h)}(x) - \sum_{\mu=1}^m \left\{ \mathfrak{G}^{(\mu)}(x, \eta_{\mu}, \lambda) \sum_{h=0}^{\tau} \lambda^{-h-1} f^{(h)}(\eta_{\mu}) \right. \\ \left. \left. + \int_{\eta_{\mu}}^x \lambda^{-\tau-1} \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \mathfrak{R}(x_1) f^{(\tau+1)}(x_1) dx_1 \right\} \right\}.$$

If in this relation the index τ is chosen to coincide with that of the formal expansions (15.10), and in these latter the evaluations (17.2) are substituted, the formulas reduce to

$$(17.3) \quad \mathfrak{g}_x^{(l)}(x) = \frac{1}{2\pi i} \int_{\Gamma_x} \sum_{h=0}^{\tau} f^{(h)}(x) \frac{d\lambda}{\lambda^{h-l+1}} \\ - \frac{1}{2\pi i} \int_{\Gamma_x} \sum_{\mu=1}^m \int_{\eta_{\mu}}^x \lambda^{l-\tau-1} \mathfrak{G}^{(\mu)}(x, x_1, \lambda) \mathfrak{R}(x_1) f^{(\tau+1)}(x_1) dx_1 d\lambda.$$

Consider the final member on the right of this relation. By the formula (16.4) its integrand consists of a finite number of terms, of which

$$(17.4) \quad \int_{\eta_k}^x \mathfrak{S}_{\beta, l, k}^{(\mu)}(x, x_1, \lambda) \frac{\lambda^{l-\tau-1} \exp [\lambda \{ \Psi_{\beta, l, k}^{(\mu)}(x) - R_k(x_1) \}]}{D(\lambda)} d\lambda$$

is a typical one. By the conditions of regularity of the boundary problem as to the chosen x , there exists in the x_1 plane a curve which may be taken as the path of integration in (17.4), and which maps under the transformation (16.7) upon a locus no point of which lies outside of the polygon P . Inasmuch as the vertices of the polygon P all lie by construction at points of the set Ω_α , it follows that whatever λ may be, there corresponds to it a choice of the index α such that the real part of $\lambda\Omega_\alpha$ at least equals the real part of the exponent in (17.4), that is, such that

$$| \exp [\lambda \{ \Psi_{\beta, l, k}^{(\mu)}(x) - R_k(x_1) \} - \lambda\Omega_\alpha] | \leq 1$$

uniformly in x_1 .

Now it was observed in §12 that the reciprocals of the expressions (12.2) for all choices of α are bounded when λ is restricted to the contours of a set Γ_κ , as may be assumed in the present discussion. The scalar factor in the integrand of (17.4) is, therefore, of the order of the $(l-\tau-\rho-1)$ th power of λ when $|\lambda| > N$. In virtue of the relation (16.7), the order of the entire integrand exceeds that by no more than the θ th power of λ , and since the result is uniform as to x_1 on the path of integration, that is true for the integral itself. Thus the final member of the relation (17.3) calls for the integral over the contour Γ_κ of a function which is of the order of λ to the power $(l+\theta-\rho-1-\tau)$. Of the integers θ, ρ, l, τ , which thus come into question, the first two are determined by the boundary problem, and the third is merely indicative of which of the expansions (15.8) is in question. The integer τ , though it is definitive for the formal expansions under consideration, has hitherto remained unspecified. Let it be chosen now as nonnegative and at least equal to the integer $(\theta-\rho+1)$, and let the larger of the numbers τ and $\tau-(\theta-\rho+1)$ be denoted by l_1 . Then for any index l such that $l \leq l_1$, the first member on the right of the relation (17.3), which is directly integrable, has the value $f^{(l)}(x)$. The second member, having an integrand of at most the order of λ^{-2} , converges to zero as $\kappa \rightarrow \infty$, by virtue of the configurations of the contours Γ_κ . Thus at the point x ,

$$(17.5) \quad \lim_{\kappa \rightarrow \infty} \mathfrak{S}_\kappa^{(l)}(x) = f^{(l)}(x), \quad l = 0, 1, 2, \dots, l_1,$$

and, since the value $l=0$ is at all events included, the formal expansion (15.8) of the chosen vector $f(x)$ itself converges to this vector.

If, by the choice of τ , the case is one in which $l_1 \geq 1$, the series obtained by the term by term differentiation of $g^{(l)}(x)$, with $l < l_1$, is found to be identical with the series $\Re(x)g^{(l+1)}(x) + \Im(x)g^{(l)}(x)$, and, since this converges, to have the value $\Re(x)f^{(l+1)}(x) + \Im(x)f^{(l)}(x)$, a value which by (15.3) reduces to $f^{(l)'}(x)$. This follows from the fact, which was observed in §13, that the residues of the Green's matrices involved in the terms of $g^{(l)}(x)$ satisfy the differential equation (7.1a). Thus every expansion (15.8) for which $l < l_1$ is differentiable term by term at the point x , and by iteration it is seen at once that the expansion for the vector $f(x)$ itself admits of term by term differentiation to the order l_1 .

Throughout the foregoing discussion it has been assumed only that the boundary problem is regular as to the point x . If it is assumed now that the problem is regular as to a connected region of the x plane, and x is taken in this region, it will be verified without difficulty that the results of the discussion are at each stage valid uniformly as to x . The convergence of the formal expansions indicated by (17.5) is thus uniform, and this applies, in particular, to the expansion of $f(x)$ itself if τ is merely chosen as the larger of the numbers 0 and $(\theta - \rho + 1)$. Because of the uniformity of the convergence, the term by term differentiability of the expansion necessarily follows in this case, and, by a reversal of the reasoning employed above, it may be inferred therefrom that, with the index τ which was fixed upon, the expansions (15.8), for all l , converge uniformly to the respective vectors $f^{(l)}(x)$.

THE UNIVERSITY OF WISCONSIN,
MADISON, WIS.

ON THE REPRESENTATION OF A FUNCTION BY CERTAIN FOURIER INTEGRALS*

BY
HARALD CRAMÉR

1. **Introduction.** Let us consider a complex-valued function $f(t)$ of the real variable t , which is *bounded for all real t and integrable in the Lebesgue sense over every finite interval*. It is proposed to investigate the conditions under which $f(t)$ admits a representation of one of the following types:

$$(F) \quad f(t) = \int_{-\infty}^{\infty} e^{itz} dF(x),$$

where $F(x)$ is real, bounded and never decreasing;

$$(G) \quad f(t) = \int_{-\infty}^{\infty} e^{itz} dG(x),$$

where $G(x)$ is of bounded variation in $(-\infty, \infty)$; and

$$(g) \quad f(t) = \int_{-\infty}^{\infty} e^{itz} g(x) dx,$$

where $g(x)$ is absolutely integrable over $(-\infty, \infty)$. The functions $G(x)$ and $g(x)$ are not necessarily real.

We shall say that a representation of one of these types *exists*, whenever $f(t)$ is represented by the corresponding expression for *almost all* real t . If, in addition, we know a priori that $f(t)$ is continuous, it readily follows from elementary properties of the above integrals that our representation holds for *all* real t .

Now let us denote by $\mu(t)$ a function which satisfies the following conditions (1) and (2):

$$(1) \quad \int_{-\infty}^{\infty} |\mu(t)| dt \text{ is finite,}$$

$$(2a) \quad \mu(t) = \int_{-\infty}^{\infty} e^{itz} m(x) dx,$$

where $m(x)$ is real and never negative, and

$$(2b) \quad \mu(0) = \int_{-\infty}^{\infty} m(x) dx = 1.$$

* Presented to the Society, February 25, 1939; received by the editors January 31, 1939.

The functions $\mu(t) = e^{-t^2/2}$, $\mu(t) = e^{-|t|}$, and

$$(3) \quad \mu(t) = \begin{cases} 1 - |t|, & |t| \leq 1, \\ 0, & |t| \geq 1, \end{cases}$$

are examples of functions satisfying these conditions. The corresponding $m(x)$ -functions are, respectively,

$$\frac{1}{(2\pi)^{1/2}} e^{-x^2/2}, \quad \frac{1}{\pi(1+x^2)}, \quad \frac{1 - \cos x}{\pi x^2}.$$

For any positive ϵ we denote by $g_\epsilon(x)$ the function defined for all real x by the absolutely convergent integral

$$(4) \quad g_\epsilon(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz} \mu(\epsilon t) f(t) dt.$$

Obviously $g_\epsilon(x)$ is bounded and everywhere continuous.

We then have for any particular $\mu(t)$ satisfying (1) and (2) the following necessary and sufficient conditions for the existence of a representation of $f(t)$ according to (F), (G), or (g):

Type (F). $g_\epsilon(x)$ should be real and never negative for $0 < \epsilon < 1$ and for all real x .

Type (G). $\int_{-\infty}^{\infty} |g_\epsilon(x)| dx < \text{const.}$ for $0 < \epsilon < 1$.

Type (g). $g_\epsilon(x)$ should satisfy the condition for type (G), and further

$$\lim_{\epsilon' \rightarrow 0} \int_{-\infty}^{\infty} |g_\epsilon(x) - g_{\epsilon'}(x)| dx = 0.$$

If a given function $f(t)$ satisfies one of these conditions for *one* particular function $\mu(t)$, it follows that the same condition is automatically satisfied for *all* $\mu(t)$ satisfying (1) and (2).

Proofs of the conditions will be given in §§3-5. In §7, it will be shown that similar conditions hold for functions $f(t_1, \dots, t_k)$ of any number of variables.

2. **A particular case.** Choosing for $\mu(t)$ the particular function given by (3), we obtain, writing $A = 1/\epsilon$,

$$(5) \quad \begin{aligned} g_\epsilon(x) &= \frac{1}{2\pi} \int_{-A}^A \left(1 - \frac{|t|}{A}\right) f(t) e^{-itz} dt \\ &= \frac{1}{2\pi A} \int_0^A \int_0^A f(t-u) e^{-iz(t-u)} dt du. \end{aligned}$$

In this particular case, our conditions are analogous to those given by

Hausdorff [4] with respect to the problem of representing a sequence of numbers c_k , ($k=0, \pm 1, \pm 2, \dots$), in the form

$$c_k = \int_0^{2\pi} e^{ikx} dF(x)$$

or in one of the similar forms corresponding to (G) or (g).

Our condition for type (F) constitutes, in the particular case when $\mu(t)$ is given by (3), a simplified form of a well known theorem due to Bochner (cf. §7). For type (G), Bochner [2] and Schoenberg [6] have given a necessary and sufficient condition which is, however, fundamentally different from ours.

Some applications of our conditions to the theory of random processes will be given in a forthcoming paper.

3. **Representation of type (F).** In the case of a representation

$$(F) \quad f(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$$

with a real, bounded, and never decreasing $F(x)$, it is almost obvious that our condition is *necessary*. We obtain, in fact, from (4)

$$\begin{aligned} g_\epsilon(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \mu(\epsilon t) dt \int_{-\infty}^{\infty} e^{itv} dF(y) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dF(y) \int_{-\infty}^{\infty} e^{-it(x-v)} \mu(\epsilon t) dt, \end{aligned}$$

the inversion of the order of integration being justified by the absolute convergence of the integrals. According to (1) and (2) we have, however, almost everywhere

$$m(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \mu(t) dt,$$

so that $g_\epsilon(x)$ is given by the "Faltung"

$$g_\epsilon(x) = \frac{1}{\epsilon} \int_{-\infty}^{\infty} m\left(\frac{x-y}{\epsilon}\right) dF(y),$$

which is obviously real and never negative.

In order to show that the condition is also *sufficient*, we consider the identity

$$\int_{-2A}^{2A} \left(1 - \frac{|x|}{2A}\right) e^{-itx} dx = 2A \left(\frac{\sin At}{At}\right)^2,$$

which holds for every $A > 0$. Multiplying by $(2\pi)^{-1}\mu(\epsilon t)f(t)dt$ and integrating with respect to t over $(-\infty, \infty)$, we obtain, according to (4),

$$\int_{-2A}^{2A} \left(1 - \frac{|x|}{2A}\right) g_{\epsilon}(x) dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \left(\frac{\sin t}{t}\right)^2 \mu(\epsilon t/A) f(t/A) dt.$$

Now $|\mu(t)| \leq 1$ by (2), and $f(t)$ is bounded by hypothesis; say $|f(t)| \leq c$. Thus if $g_{\epsilon}(x)$ is real and never negative, we conclude

$$\int_{-2A}^{2A} \left(1 - \frac{|x|}{2A}\right) g_{\epsilon}(x) dx \leq c$$

for $0 < \epsilon < 1$ and for all positive A . This obviously implies

$$(6) \quad \int_{-\infty}^{\infty} g_{\epsilon}(x) dx \leq c$$

for $0 < \epsilon < 1$.

From (4) and (6) we then obtain for almost all values of x

$$(7) \quad \mu(\epsilon t)f(t) = \int_{-\infty}^{\infty} e^{itz} g_{\epsilon}(x) dx.$$

Now, since both $\mu(t)$ and the integral are continuous functions of t , it follows that it is possible to find a continuous function $f^*(t)$ which coincides with $f(t)$ for almost all real t . We then have

$$\mu(\epsilon t)f^*(t) = \int_{-\infty}^{\infty} e^{itz} g_{\epsilon}(x) dx$$

for all real t and for $0 < \epsilon < 1$.

Consider now the last relation for a sequence of values of ϵ tending to zero. As $\mu(0) = 1$, the left-hand side tends to $f^*(t)$ uniformly in every finite interval. According to a fundamental theorem on characteristic functions due to Lévy [5] (cf. also Bochner [1]), we then have for all real t

$$f^*(t) = \int_{-\infty}^{\infty} e^{itz} dF(x)$$

where $F(x)$ is real and never decreasing. As $f^*(t) = f(t)$ for almost all t , this proves our assertion.

4. Representation of type (G).† If we have for almost all t

$$(G) \quad f(t) = \int_{-\infty}^{\infty} e^{itz} dG(x),$$

† The author is indebted to Mr. E. Frithiofson of Lund for a remark leading to a simplification of the condition for this type.

where $G(x)$ is of bounded variation in $(-\infty, \infty)$, we obtain as in the preceding paragraph

$$g_\epsilon(x) = \frac{1}{\epsilon} \int_{-\infty}^{\infty} m\left(\frac{x-y}{\epsilon}\right) dG(y)$$

and thus, $m(x)$ being never negative,

$$(8) \quad \int_{x_1}^{x_2} |g_\epsilon(x)| dx \leq \int_{-\infty}^{\infty} |dG(y)| \int_{(x_1-y)/\epsilon}^{(x_2-y)/\epsilon} m(x) dx.$$

Hence we obtain, using (2b),

$$\int_{-\infty}^{\infty} |g_\epsilon(x)| dx \leq \int_{-\infty}^{\infty} |dG(y)|$$

for $0 < \epsilon < 1$. Thus our condition is *necessary*.

In order to show that the condition is also *sufficient* we observe that, owing to the convergence of $\int_{-\infty}^{\infty} |g_\epsilon(x)| dx$, the relation (4) may be converted into (7) for almost all real t . As in the preceding section, it follows that there is a continuous function $f^*(t)$ which coincides with $f(t)$ for almost all real t . We then have as before

$$\mu(\epsilon t) f^*(t) = \int_{-\infty}^{\infty} e^{itz} g_\epsilon(x) dx$$

for all real t and for $0 < \epsilon < 1$. Putting

$$G_\epsilon(x) = \int_{-\infty}^x g_\epsilon(y) dy,$$

we may write this as

$$\mu(\epsilon t) f^*(t) = \int_{-\infty}^{\infty} e^{itz} dG_\epsilon(x).$$

When ϵ tends to zero, the left-hand side of this relation tends to $f^*(t)$ uniformly in every finite interval. On the other hand, $\int_{-\infty}^{\infty} |g_\epsilon(x)| dx$ is uniformly bounded for $0 < \epsilon < 1$, so that $G_\epsilon(x)$ is of uniformly bounded variation in $(-\infty, \infty)$. It is well known that we can always find a sequence $\epsilon_1, \epsilon_2, \dots$ tending to zero and a function $G(x)$ of bounded variation in $(-\infty, \infty)$ such that

$$(9) \quad G(x) = \lim_{n \rightarrow \infty} G_{\epsilon_n}(x) = \lim_{n \rightarrow \infty} \int_{-\infty}^x g_{\epsilon_n}(y) dy$$

in every point of continuity x of $G(x)$. It then follows from a lemma given by

Bochner [2, p. 274], that we have for all real t

$$f^*(t) = \int_{-\infty}^{\infty} e^{itz} dG(x).$$

As $f^*(t) = f(t)$ for almost all t , this proves our assertion.

For a later purpose it will now be shown that, if our condition for type (G) is satisfied, then the integral

$$(10) \quad \int_{-\infty}^{\infty} |g_{\epsilon}(x)| dx$$

is *uniformly convergent* for $0 < \epsilon < 1$. If the condition is satisfied, we already know that $f(t)$ admits a representation of type (G). Now let $\delta > 0$ be given. We can then choose $y_0 > 0$ and $x_0 > y_0$ such that

$$\int_{y_0}^{\infty} |dG(y)| < \delta, \quad \int_{x_0-y_0}^{\infty} m(x) dx < \delta.$$

Obviously x_0 and y_0 can be chosen independently of ϵ . For $x_2 > x_1 > x_0$ and for $0 < \epsilon < 1$, we then conclude from (8) and (2b) that

$$\begin{aligned} \int_{x_1}^{x_2} |g_{\epsilon}(x)| dx &< \delta \int_{-\infty}^{y_0} |dG(y)| + \int_{y_0}^{\infty} |dG(y)| \\ &< \delta \left[1 + \int_{-\infty}^{\infty} |dG(y)| \right]. \end{aligned}$$

A similar inequality evidently holds for negative values of x_1 and x_2 , and thus the uniform convergence of (10) is established.

5. Representation of type (g). As in the preceding cases, we begin by proving that our condition is *necessary*. Any representation of type (g) being a particular case of type (G), it is obvious that the first part of the condition is necessary. It thus remains to show that, if

$$(g) \quad f(t) = \int_{-\infty}^{\infty} e^{itz} g(x) dx$$

holds for almost all t , where $g(t)$ is absolutely integrable over $(-\infty, \infty)$, then

$$\lim_{\substack{\epsilon \rightarrow 0 \\ \epsilon' \rightarrow 0}} \int_{-\infty}^{\infty} |g_{\epsilon}(x) - g_{\epsilon'}(x)| dx = 0.$$

As $|g_{\epsilon} - g_{\epsilon'}| \leq |g - g_{\epsilon}| + |g - g_{\epsilon'}|$, it is only necessary to prove that

$$(11) \quad \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} |g(x) - g_{\epsilon}(x)| dx = 0.$$

According to the preceding section, it follows from the first part of the condition that the integral (10) converges uniformly for $0 < \epsilon < 1$. Given $\delta > 0$, we can thus choose $x_0 = x_0(\delta)$ such that

$$(12) \quad \int_{|x| > x_0} |g(x) - g_\epsilon(x)| dx < \delta$$

for $0 < \epsilon < 1$.

We now choose a function $g^*(x)$, *bounded and continuous* for all real x , such that

$$(13) \quad \int_{-\infty}^{\infty} |g(x) - g^*(x)| dx < \delta$$

with

$$|g^*(x)| < K = K(\delta),$$

and we put

$$g_\epsilon^*(x) = \frac{1}{\epsilon} \int_{-\infty}^{\infty} m\left(\frac{x-y}{\epsilon}\right) g^*(y) dy.$$

We then have

$$(14) \quad \begin{aligned} \int_{-x_0}^{x_0} |g(x) - g_\epsilon(x)| dx &\leq \int_{-x_0}^{x_0} |g(x) - g^*(x)| dx + \int_{-x_0}^{x_0} |g^*(x) - g_\epsilon^*(x)| dx \\ &\quad + \int_{-x_0}^{x_0} |g_\epsilon^*(x) - g_\epsilon(x)| dx. \end{aligned}$$

According to (13), the first term on the right-hand side is less than δ . We further have, using (2b),

$$(15) \quad g^*(x) - g_\epsilon^*(x) = \frac{1}{\epsilon} \int_{-\infty}^{\infty} m\left(\frac{x-y}{\epsilon}\right) (g^*(x) - g^*(y)) dy.$$

Now, $g^*(x)$ is uniformly continuous in every finite interval. The numbers δ and x_0 being given, we can thus choose $h = h(\delta, x_0)$ such that for $|x| < x_0$, $|x-y| < h$ we have

$$|g^*(x) - g^*(y)| < \delta/x_0.$$

We can further choose $y_0 = y_0(\delta, x_0, K)$ such that

$$\int_{|y| > y_0} m(y) dy < \frac{\delta}{2Kx_0}.$$

For any ϵ such that $0 < \epsilon < h/y_0$, we then obtain from (15)

$$|g^*(x) - g_\epsilon^*(x)| < \frac{\delta}{x_0} + 2K \int_{|y|>h/\epsilon} m(y) dy < \frac{2\delta}{x_0}$$

and

$$(16) \quad \int_{-x_0}^{x_0} |g^*(x) - g_\epsilon^*(x)| dx < 4\delta.$$

Finally, we have

$$g_\epsilon^*(x) - g_\epsilon(x) = \frac{1}{\epsilon} \int_{-\infty}^{\infty} m\left(\frac{x-y}{\epsilon}\right) (g^*(y) - g(y)) dy,$$

and hence by (13)

$$(17) \quad \begin{aligned} \int_{-x_0}^{x_0} |g_\epsilon^*(x) - g_\epsilon(x)| dx &\leq \int_{-\infty}^{\infty} |g^*(y) - g(y)| dy \int_{(x_0-y)/\epsilon}^{(x_0+y)/\epsilon} m(x) dx \\ &\leq \int_{-\infty}^{\infty} |g^*(y) - g(y)| dy < \delta. \end{aligned}$$

From (12), (14), (16), and (17) we then obtain

$$\int_{-\infty}^{\infty} |g(x) - g_\epsilon(x)| dx < 7\delta$$

for all sufficiently small $\epsilon > 0$, so that (11) is proved.

We now have to show that our condition is *sufficient*. From the first part of the condition, it follows by the preceding paragraph that we have for almost all real t

$$(G) \quad f(t) = \int_{-\infty}^{\infty} e^{itz} dG(x),$$

where $G(x)$ is of bounded variation in $(-\infty, \infty)$, and according to (9)

$$(18) \quad G(x) = \lim_{n \rightarrow \infty} G_n(x) = \lim_{n \rightarrow \infty} \int_{-\infty}^x g_n(y) dy$$

in every point of continuity x of $G(x)$.

From the second part of the condition it follows, however, that there is a function $g(x)$, absolutely integrable over $(-\infty, \infty)$, such that

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} |g(y) - g_n(y)| dy = 0.$$

Hence we obtain for all real x

$$\lim_{n \rightarrow \infty} G_n(x) = \lim_{n \rightarrow \infty} \int_{-\infty}^x g_n(y) dy = \int_{-\infty}^x g(y) dy.$$

It then finally follows from (18) that

$$G(x) = \int_{-\infty}^x g(y) dy$$

for almost all x , and

$$f(t) = \int_{-\infty}^{\infty} e^{itx} g(x) dx$$

for almost all t , so that the proof is completed.

If the first part of our condition for type (g) is replaced by the condition given above for type (F), it is readily seen that we obtain a necessary and sufficient condition for representation of type (g) with a real and non-negative $g(x)$.

It may be worth while to point out that the first part of our condition for type (g) is *not* contained in the second part. This is shown by the example

$$f(t) = \begin{cases} i(-1-t), & -1 < t < 0, \\ i(1-t), & 0 < t < 1, \\ 0, & t = 0, \quad |t| \geq 1. \end{cases}$$

In the particular case when $\mu(t)$ is given by (3), this function yields for $0 < \epsilon < 1$

$$g_\epsilon(x) = \frac{x - \sin x}{\pi x^2} + \epsilon \frac{x \sin x - 2(1 - \cos x)}{\pi x^3},$$

so that the second part of the condition is satisfied but not the first part. Accordingly, no representation of any of our three types exists, which is also directly seen from the behaviour of $f(t)$ near $t=0$.

6. The case of an unbounded $f(t)$. In all the preceding paragraphs it has been a priori assumed that $f(t)$ is *bounded*. It will, however, be seen that this assumption has only been used on two occasions; namely (a) to ensure the absolute convergence of the integral (4) which defines $g_\epsilon(x)$, and (b) for the proof that our condition for type (F) is sufficient.

Let us now omit this assumption and consider the class of all functions $f(t)$ which are integrable over any finite interval. Let us further choose for $\mu(t)$ the particular function given by (3). As this function is equal to zero for $|t| \geq 1$, it is obvious that the integral (4) will still be absolutely convergent for any positive ϵ .

Thus the conditions for types (G) and (g) remain true under the present con-

ditions, while in the condition for type (F) it will have to be explicitly stated that $|f(t)|$ should be less than a constant K for almost all values of t .

The necessity of this addition to the condition for type (F) is shown by the example $f(t) = |t|^{-\alpha}$, ($0 < \alpha < 1$), where obviously no K can be found such that $|f(t)| < K$ for almost all t . The corresponding function $g_\epsilon(x)$ can be shown to be positive for $0 < \epsilon < 1$ and for all real x , although evidently no representation of type (F) exists.

7. Functions of several variables. So far we have only considered functions $f(t)$ of a single variable t . All our considerations can, however, be extended to functions $f(t_1, \dots, t_k)$ of any finite number of real variables. This requires only a straightforward generalization of our above arguments, based on the elementary properties of Fourier integrals in several variables. The only delicate point arising in this connection is the generalization to several variables of Bochner's lemma used in the proof of our condition for type (G). This generalization is, however, easily performed by means of a general induction method due to Cramér and Wold (Cramér [3, p. 104]).

We obtain in this way direct generalizations of our above conditions, the auxiliary functions $\mu(t)$ and $g_\epsilon(x)$ being replaced by the functions of k variables obtained when, in (1), (2), and (4), we regard x , t , and ϵt as abbreviations for (x_1, \dots, x_k) , (t_1, \dots, t_k) , and $(\epsilon t_1, \dots, \epsilon t_k)$, respectively, and put $tx = t_1x_1 + \dots + t_kx_k$, the integrals being taken over the k -dimensional euclidean space R_k . Moreover, in the definition (4) of $g_\epsilon(x)$ the factor $1/2\pi$ has to be replaced by $1/(2\pi)^k$.

For $\mu(t_1, \dots, t_k)$ we may, for example, choose any function of the form $\mu(t_1)\mu(t_2) \dots \mu(t_k)$, where $\mu(t)$ satisfies the conditions (1) and (2). The definition (4) of $g_\epsilon(x)$ will then be replaced by

$$(19) \quad g_\epsilon(x_1, \dots, x_k) = \frac{1}{(2\pi)^k} \int_{R_k} e^{-i(t_1x_1 + \dots + t_kx_k)} \mu(\epsilon t_1) \dots \mu(\epsilon t_k) f(t_1, \dots, t_k) dt_1 \dots dt_k.$$

In particular, we obtain in this way the following new characterization of the class of *positive definite functions* of k variables as defined by Bochner [1, p. 406]. Bochner has established the identity of this class with the class of functions represented for all real t_r by the expression

$$f(t_1, \dots, t_k) = \int_{R_k} e^{i(t_1x_1 + \dots + t_kx_k)} dF(x_1, \dots, x_k),$$

where F is real, bounded, and, for each particular x_r , never decreasing. The class of positive definite functions such that $f(0, \dots, 0) = 1$ is thus identical

with the class of characteristic functions of k -dimensional random variables in the sense of the theory of probability (cf. Cramér [3]). Using our generalized condition for type (F) we then conclude:

For any particular $\mu(t)$ satisfying (1) and (2) a necessary and sufficient condition that a given bounded and continuous function $f(t_1, \dots, t_k)$ should be positive definite is that $g_\epsilon(x_1, \dots, x_k)$ as defined by (19) should be real and never negative for $0 < \epsilon < 1$ and for all real x_1, \dots, x_k .

Choosing, in particular, in (19) the special function $\mu(t)$ given by (3), we obtain in analogy with (5), writing $A = 1/\epsilon$,

$$\begin{aligned} g_\epsilon(x_1, \dots, x_k) &= \frac{1}{(2\pi)^k} \int_{-A}^A \cdots \int_{-A}^A f(t_1, \dots, t_k) \\ &\quad \cdot \prod_{r=1}^k \left[\left(1 - \frac{|t_r|}{A} \right) e^{-it_r x_r} \right] dt_1 \cdots dt_k \\ &= \frac{1}{(2\pi A)^k} \int_0^A \cdots \int_0^A f(t_1 - u_1, \dots, t_k - u_k) \\ &\quad \cdot \exp \left(-i \sum_1^k x_r t_r \right) \cdot \exp \left(i \sum_1^k x_r u_r \right) dt_1 \cdots dt_k du_1 \cdots du_k. \end{aligned}$$

Now Bochner's original condition for a positive definite function requires that

$$\int_a^A \cdots \int_a^A f(t_1 - u_1, \dots, t_k - u_k) \rho(t_1, \dots, t_k) \cdot \overline{\rho(u_1, \dots, u_k)} dt_1 \cdots dt_k du_1 \cdots du_k \geq 0$$

for all real a, A and for all continuous functions $\rho(t_1, \dots, t_k)$. Thus our condition, with the particular choice of $\mu(t)$ according to (3), involves a considerable simplification.

REFERENCES

1. S. Bochner, *Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse*, Mathematische Annalen, vol. 108 (1933), p. 378.
2. ———, *A theorem on Fourier-Stieltjes integrals*, Bulletin of the American Mathematical Society, vol. 40 (1934), p. 271.
3. H. Cramér, *Random Variables and Probability Distributions*, Cambridge, 1937.
4. F. Hausdorff, *Momentprobleme für ein endliches Intervall*, Mathematische Zeitschrift, vol. 16 (1923), p. 220.
5. P. Lévy, *Calcul des Probabilités*, Paris, 1925.
6. I. J. Schoenberg, *Remark on the preceding note by Bochner*, Bulletin of the American Mathematical Society, vol. 40 (1934), p. 277.

UNIVERSITY OF STOCKHOLM,
STOCKHOLM, SWEDEN

GENERAL THEORY OF SINGULAR INTEGRAL EQUATIONS WITH REAL KERNELS*

BY

W. J. TRJITZINSKY

1. **Introduction.** Amongst the outstanding theories of integral equations of particular importance from our present point of view are those due to Vito Volterra,† I. Fredholm,‡ D. Hilbert,§ E. Schmidt,|| and T. Carleman.¶ With respect to generality these contributions, in the order mentioned, form an ascending hierarchy of theories, with those of Hilbert and Schmidt essentially on the same level, while the developments of Carleman present the culminating aspects. In considering integral equations of the form

$$(1.1) \quad \phi(x) - \lambda \int_a^b K(x, y)\phi(y)dy = f(x),$$

$$(1.2) \quad \phi(x) - \lambda \int_a^b K(x, y)\phi(y)dy = 0$$

[$f(x)$ given on (a, b) ; real $K(x, y)$ given on $a \leq x, y \leq b$].

one may, with advantage and without any substantial loss of generality, confine oneself to symmetric kernels $K(x, y)$,

$$K(x, y) = K(y, x).$$

This can be inferred on the basis of certain considerations of Pèrès.**

In the sequel, unless the contrary is stated, all kernels involved will be supposed symmetric. All integrals not in the sense of Stieltjes will be in the sense of Lebesgue.

Whenever††

* Presented to the Society, September 7, 1939; received by the editors March 7, 1939.

† An exposition of Volterra's work and of many other developments in the field of integral equations as well as an extensive bibliography can be found in the book by V. Volterra and J. Pèrès, *Théorie Générale des Fonctionnelles*, vol. 1, Paris, 1936.

‡ Cf. reference on page 344 of Volterra and Pèrès, loc. cit.

§ D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, Leipzig and Berlin, 1912.

|| Cf. reference on page 347 of Volterra and Pèrès, loc. cit.

¶ T. Carleman, *Sur les Équations Intégrales Singulières à Noyau Réel et Symétrique*, Uppsala, 1923; T. Carleman, *La théorie des équations intégrales singulières et les applications*, Annales de l'Institut H. Poincaré, vol. 1, pp. 401-430.

** Cf. the book of Volterra and Pèrès, loc. cit., pp. 305-306 and pp. 263-264.

†† That is, the integrals $\int_a^b \int_a^b K^2(x, y)dx dy$, $\int_a^b f^2(x)dx$ exist.

$$(1.3) \quad K(x, y) \in L_2 \text{ (in } x, y), \quad f(x) \in L_2,$$

the essential results of the Fredholm theory will hold.*

The results of Hilbert's theory will hold in the essential particulars if

$$(1.4) \quad K(x, y) \in L_2 \quad (\text{in } y; \text{ for almost all } x),$$

$$(1.4a) \quad \int_a^b \int_a^b K(x, y) \phi(x) \phi(y) dx dy \leq k^2 \int_a^b \phi^2(x) dx$$

(k independent of $\phi(x)$).

The highly important investigations of Carleman extend these theories as follows. In some of his investigations (1.4) is assumed (for all x except for $x = \xi_1, \xi_2, \dots$; the ξ , possessing merely a finite number of limiting points), while condition (1.4a) is deleted; in certain other developments he retains (1.4), deletes (1.4a) and assumes the mean continuity relation

$$(1.4b) \quad \lim_{x_1 \rightarrow x_2} \int_a^b [K(x_1, y) - K(x_2, y)]^2 dy = 0 \quad (x_1, x_2 \neq \xi_1, \xi_2, \dots).$$

Carleman also has a still more general theory in which the conditions (1.4), (1.4a), (1.4b) are deleted and it is merely assumed that $K(x, y)$ is a limit (in the ordinary sense or in the mean square with respect to y) of kernels satisfying (1.3).

The applications of Carleman's results (or of suitable extensions of them) have been numerous and important; witness, for instance, the application to the Schrödinger wave equation† and to nonlinear ordinary differential equations (of the type occurring in dynamics).‡

Our object in the present work is to develop a theory of equations (1.1) (with $f(x) \in L_2$), (1.2) with kernels $K(x, y)$ which, while not necessarily of Carleman's type, are limits (in one sense or another) of kernels of Carleman's type. The kernels of this description will be said to be of rank two. More generally we shall develop theories of equations whose kernels $K(x, y)$ are of any rank n (≥ 2). In this connection $K(x, y)$ will be said to be of rank n if $K(x, y)$, while not necessarily of rank $n-1$, is a limit (in a suitable sense)§ of kernels of ranks less than n . In accordance with the above, Carleman's kernels are said to be of rank 1.

* A more precise statement in this regard can be found in Carleman, *Annales de l'Institut H. Poincaré*, loc. cit., pp. 401-402.

† T. Carleman, *Sur la théorie mathématique de l'équation de Schrödinger*, *Arkiv för Matematik, Astronomi och Fysik*, vol. 24B (1934), pp. 1-7.

‡ T. Carleman, *Application de la théorie des équations intégrales linéaires aux systèmes d'équations différentielles non linéaires*, *Acta Mathematica*, vol. 59, pp. 63-87.

§ More precise formulation will be given in the sequel.

In these pages we shall consider also equations whose kernels are of transfinite rank.

In the sequel Carleman's book will be referred to as (C).

We shall have occasion to use the following known theorems.

THEOREM 1.1. (Helly.) *Let $\alpha(x, n)$ ($n = 1, 2, \dots$) be of bounded variation for $a \leq x \leq b$. If $\text{Var. } \alpha(x, n) < A$ ($n = 1, 2, \dots$; A independent of n) and if $\lim_n \alpha(x, n) = \alpha(x)$, then*

$$\lim_n \int_a^b w(x) d_x \alpha(x, n) = \int_a^b w(x) d_x \alpha(x) \quad (\text{for } w(x) \text{ continuous}).$$

THEOREM 1.2. (F. Riesz.) *Suppose $f_\nu(x) \in L_2$, $g_\nu(x) \in L_2$ (for $\nu = 1, 2, \dots$ and x on (a, b)) and $f_\nu(x) \rightarrow f(x)$, $g_\nu(x) \rightarrow g(x)$ (almost everywhere). Then, provided*

$$\int_a^b g_\nu^2(x) dx < c, \quad |f_\nu(x)| < \gamma(x) \in L_2 \quad (\nu = 1, 2, \dots),$$

one has

$$\lim_n \int_a^b f_n(x) g_n(x) dx = \int_a^b f(x) g(x) dx.$$

THEOREM 1.3. (F. Riesz.) *If $f_\nu(x) \in L_2$ on (a, b) ($\nu = 1, 2, \dots$) and if*

$$\int_a^b f_\nu^2(x) dx < M \quad (\nu = 1, 2, \dots),$$

then there exists a subsequence $\{f_{\nu_j}(x)\}$ ($\nu_1 < \nu_2 < \dots$) such that, as $j \rightarrow \infty$, $f_{\nu_j}(x) \rightarrow f(x)$ weakly; that is,

$$\lim_j \int_a^x f_{\nu_j}(x) dx = \int_a^x f(x) dx;$$

moreover $\int_a^b f^2(x) dx \leq M$.

THEOREM 1.4. (F. Riesz.) *Let $f_\nu(x) \in L_2$ on (a, b) ($\nu = 1, 2, \dots$) and suppose $f_\nu(x) \rightarrow f(x)$ weakly; then, provided $g(x) \in L_2$, one has*

$$\lim_\nu \int_a^b f_\nu(x) g(x) dx = \int_a^b f(x) g(x) dx.$$

THEOREM 1.5. (T. Carleman.*) *If $\int_a^b f^2(x) dx < c$, $f_\nu(x) \rightarrow f(x)$ weakly, $g_n(x) \rightarrow g(x)$ and $|g_n(x)| < \gamma(x) \in L_2$, then*

$$\lim_n \int_a^b f_n(x) g_n(x) dx = \int_a^b f(x) g(x) dx.$$

* (C), pp. 132-133.

Another theorem necessary for our purposes will be the theorem of (C, pp. 21, 22), which constitutes an extension by Carleman of a result due to Hilbert. This theorem, in the sequel referred to as the "Compactness Theorem," gives conditions under which there exists a sequence of values δ_r ($r=1, 2, \dots; \delta_r \rightarrow 0$) such that

$$\lim_{\delta_r \rightarrow 0} f(\lambda, x_1, \dots, x_n | \delta_r) = F(\lambda, x_1, \dots, x_n),$$

where $f(\lambda, x_1, \dots, x_n | \delta)$ is a given family of functions defined for (x_1, \dots, x_n) in a domain D for every λ on (α, β) . On account of the length of this theorem the reader will be merely referred to (C, pp. 21, 22).

In the sequel we shall give examples of kernels which come under our classification and which at the same time are not of Carleman's type.

In §2 (Definition 2.2) will be introduced kernels of finite rank n belonging to classes designated as H_n . The main results for $K(x, y) \in H_n$ are given in Theorems 4.1, 5.1, 7.1.

In §10 (Definition 10.1) will be specified kernels of transfinite ranks β (β of the second class), the results for which will be given in Theorems 11.1, 11.2.

2. Kernels of class H_n . Let

$$(2.1) \quad E = E^0 = (I_1, I_2, \dots)$$

be a denumerable set of points on the closed interval (a, b) . Let us take E reducible closed with, let us say, the n th derived set,

$$(2.1a) \quad E^n = (I_1^n, I_2^n, \dots) = (s_1, s_2, \dots, s_k)$$

consisting of a finite number of points (with at least one point present). The 1st, 2d, \dots , $(n-1)$ st derived sets of E will then be denumerable sets

$$(2.1b) \quad E^v = (I_1^v, I_2^v, \dots) \quad (v = 1, 2, \dots, n-1),$$

each actually containing an infinity of points.

DEFINITION 2.1. A set E , given by (2.1) and satisfying the above conditions, will be said to belong to R_n , $E \in R_n$.*

Given a set $E \in R_{n-1}$ ($n \geq 1$), we shall form sets of closed intervals

$$\Delta^0(\delta_0), \Delta^1(\delta_1), \dots, \Delta^{n-1}(\delta_{n-1})$$

as follows. The intervals of $\Delta^0(\delta_0)$ will be

$$(2.2) \quad \Delta^0(\delta_0) = (s_v - \delta_0, s_v + \delta_0) \quad [v = 1, 2, \dots, k; E^{n-1} = (s_1, \dots, s_k)].$$

Here and in the sequel the parts of the intervals exterior to (a, b) will be discarded.

* If $E \in R_0$, E consists of a finite number of points.

In (2.2) $\delta_0 (>0)$ will be chosen sufficiently small so that no two intervals of (2.2) will have points in common; moreover, δ_0 is to be taken so that *no end point of the $\Delta^0(\delta_0)$ should be coincident with a point of E* (except, perhaps, a or b ; analogous statements are implied in the sequel).*

With $\delta_0 (>0)$ chosen as stated above, consider the set

$$(2.2a) \quad \Gamma(\delta_0) = (a, b) - \sum_{\nu=1}^k \Delta^0(\delta_0).$$

It is open. Hence, since the limiting points s_ν ($\nu=1, \dots, k$) of E^{n-2} are all in the intervals (2.2), as specified, we observe that, on one hand, there is only a finite number of points of E^{n-2} , let us say

$$(2.2b) \quad s_2^{n-2}, s_2^{n-2}, \dots, s_{m(\delta_0)}^{n-2} \quad (m(\delta_0) \rightarrow \infty, \text{ as } \delta_0 \rightarrow 0),$$

in $\Gamma(\delta_0)$ and that, on the other hand, these points (2.2b) can be enclosed in closed intervals (whose totality constitutes the set $\Delta^1(\delta_1)$)

$$(2.3) \quad \Delta^1(\delta_1) = (s_\nu^{n-2} - \delta_1, s_\nu^{n-2} + \delta_1) \quad (\nu = 1, 2, \dots, m(\delta_0))$$

so that with $\delta_1 (>0)$ sufficiently small and suitably chosen the following will be true. The intervals

$$(2.3a) \quad \Delta^0(\delta_0) \quad (\nu = 1, \dots, k), \quad \Delta^1(\delta_1) \quad (\nu = 1, \dots, m(\delta_0))$$

are all without common points; moreover, *no end point of any interval $\Delta^1(\delta_1)$ is coincident with a point of E* . It is to be noted that the intervals (2.3a) will certainly be without common points if we take $\delta_1 \leq \delta_1(\delta_0)$, where $\delta_1(\delta_0) (>0)$ is sufficiently small but, generally, depends on δ_0 .

Suppose δ_0, δ_1 chosen as stated above. The set

$$(2.4) \quad \Gamma(\delta_0, \delta_1) = (a, b) - \sum_{\nu=1}^k \Delta^0(\delta_0) - \sum_{\nu=1}^{m(\delta_0)} \Delta^1(\delta_1)$$

will be open. An infinity of points of E^{n-2} are in the intervals (2.2); all the other points of E^{n-2} —the points (2.2b)—are in the intervals (2.3); thus, all the limiting points of E^{n-3} are interior points of the closed intervals (2.3a). Consequently there is only a finite number of points of E^{n-3} , say

$$(2.4a) \quad s_1^{n-3}, s_2^{n-3}, \dots, s_{m(\delta_0, \delta_1)}^{n-3},$$

in the set $\Gamma(\delta_0, \delta_1)$ (2.4). The points (2.4a) can be enclosed in a set $\Delta^2(\delta_2)$ of closed intervals

$$(2.4b) \quad \Delta^2(\delta_2) = (s_\nu^{n-3} - \delta_2, s_\nu^{n-3} + \delta_2) \quad (\nu = 1, 2, \dots, m(\delta_0, \delta_1)).$$

* The point a (or b) will be considered interior to any subinterval (a, a') (or (b', b)) of (a, b) .

Taking $0 < \delta_2 \leq \delta_2(\delta_0, \delta_1)$ [$\delta_2(\delta_0, \delta_1)$ sufficiently small], with suitable choice of δ_2 we secure the following. The intervals

$$(2.4c) \quad \begin{aligned} \Delta^0(\delta_0) \quad (\nu = 1, \dots, k), \quad \Delta^1(\delta_1) \quad (\nu = 1, \dots, m(\delta_0)), \\ \Delta^2(\delta_2) \quad (\nu = 1, \dots, m(\delta_0, \delta_1)) \end{aligned}$$

are all without common points; no end point of these intervals is coincident with a point of E .

We continue this process a finite number of times, finally constructing the n sets of closed intervals

$$(2.5) \quad \Delta^i(\delta_i) \quad (i = 0, 1, \dots, n-1)$$

possessing properties of the following description.

The set $\Delta^i(\delta_i)$ consists of the intervals

$$(2.6) \quad \Delta^i(\delta_i) = (s_i^{n-1-i} - \delta_i, s_i^{n-1-i} + \delta_i) \quad [\nu = 1, 2, \dots, m(\delta_0, \dots, \delta_{i-1})]$$

for $i = 1, 2, \dots, n-1$. The set $\Delta^0(\delta_0)$ consists of the intervals (2.2). The numbers δ_i ,

$$(2.6a) \quad \begin{aligned} 0 < \delta_i \leq \delta_i(\delta_0, \delta_1, \dots, \delta_{i-1}) \quad (i = 1, \dots, n-1; 0 < \delta_0 \leq \delta^0; \\ \delta_i(\delta_0, \delta_1, \dots, \delta_{i-1}), \delta^0 \text{ sufficiently small}), \end{aligned}$$

are so chosen that *no point of E ((2.1)) is coincident with an end point of any of the intervals of the sets (2.5) and that all the intervals of the sets (2.5) are without common points.* The set

$$(2.6b) \quad \begin{aligned} \Gamma(\delta_0, \delta_1, \dots, \delta_{n-1}) = (a, b) - \sum_{\nu=1}^k \Delta^0(\delta_0) - \sum_{\nu=1}^{m(\delta_0)} \Delta^1(\delta_1) \\ - \sum_{\nu=1}^{m(\delta_0, \delta_1)} \Delta^2(\delta_2) - \dots - \sum_{\nu=1}^{m^1} \Delta^{n-1}(\delta_{n-1}) \end{aligned}$$

$[m^1 = m(\delta_0, \delta_1, \dots, \delta_{n-2})]$ is open and contains no points of E . The totality of all limiting points of E^{n-2} (that is, E^{n-1}) consists of the centers of the intervals of $\Delta^0(\delta_0)$. The limiting points of E^{n-3} are partly contained in $\Delta^0(\delta_0)$ and the rest of them, the points (2.2b), are centers of the intervals $\Delta^1(\delta_1)$. An infinity of limiting points of E^{n-4} (that is, the points of E^{n-3}) are in

$$\Delta^0(\delta_0) + \Delta^1(\delta_1);$$

the rest of these limiting points, the points (2.4a), are the centers of the intervals of $\Delta^2(\delta_2)$. In general, an infinity of limiting points of E^{i-1} (that is, points of E^i) are interior to the set

$$(2.7) \quad \Delta^0(\delta_0) + \Delta^1(\delta_1) + \dots + \Delta^{n-i-2}(\delta_{n-i-2});$$

the rest of the limiting points of E^{i-1} , the points

$$(2.7a) \quad s_1^i, s_2^i, \dots, s_{m^i}^i \quad (m^i = m(\delta_0, \dots, \delta_{n-i-2})),$$

consist of the centers of the intervals of the set $\Delta^{n-i-1}(\delta_{n-i-1})$. In particular, the limiting points of $E = E^0$ (2.1), that is, the points E^1 , are distributed as follows. A finite number of points of E^1 ,

$$(2.8) \quad s_1^1, s_2^1, \dots, s_{m^1}^1 \quad (m^1 = m(\delta_0, \dots, \delta_{n-3})),$$

constitute the centers of the intervals of the set $\Delta^{n-2}(\delta_{n-2})$ (2.6); all the other points of E^1 are interior to

$$(2.8a) \quad \Delta^0(\delta_0) + \dots + \Delta^{n-3}(\delta_{n-3}).$$

In the open set

$$\Gamma(\delta_0, \dots, \delta_{n-2}) = (a, b) - \Delta^0(\delta_0) - \dots - \Delta^{n-2}(\delta_{n-2})$$

there is only a finite number of points of E , say

$$(2.9) \quad s_1^0, s_2^0, \dots, s_{m^0}^0 \quad (m^0 = m(\delta_0, \dots, \delta_{n-2})).$$

These points (2.9) constitute the centers of the intervals of $\Delta^{n-1}(\delta_{n-1})$.

DEFINITION 2.2. Let E be a closed reducible set on the interval (a, b) . Suppose $E \subset R_{n-1}$ where R_{n-1} is specified by Definition (2.1). Form sets $\Delta^i(\delta_i)$ ($i=0, 1, \dots, n-1$) of closed intervals (2.6), without common points and covering the set E , as described in the text above in connection with (2.2)–(2.9).

We shall say that a real symmetric kernel $K(x, y) \in H_n$, if

$$(2.10) \quad K^{\delta_0, \delta_1, \dots, \delta_{n-1}}(x, y) \in L_2 \quad (\text{in } x, y; \text{ for } a \leq x, y \leq b),$$

the function in the first member of (2.10) being defined as follows:

$$(2.11) \quad K^{\delta_0, \dots, \delta_{n-1}}(x, y) = 0 \quad [x \text{ in } \Delta^0(\delta_0) + \Delta^1(\delta_1) + \dots + \Delta^{n-1}(\delta_{n-1}), \\ \text{while } a \leq y \leq b \text{ (or } a \leq y < x)];$$

$$(2.11a) \quad K^{\delta_0, \dots, \delta_{n-1}}(x, y) = 0 \quad [y \text{ in } \Delta^0(\delta_0) + \Delta^1(\delta_1) + \dots + \Delta^{n-1}(\delta_{n-1}), \\ \text{while } a \leq x \leq b \text{ (or } a \leq x < y)];$$

$$(2.11b) \quad K^{\delta_0, \dots, \delta_{n-1}}(x, y) = K(x, y) \quad [\text{at all other points of } a \leq x, y \leq b].$$

Moreover, this definition will be applied only if (2.10) holds as stated for all admissible* positive values δ_i ($i=0, \dots, n-1$) no matter how small.

In conformity with this definition, $K(x, y) \in H_0$ is to imply that

$$K(x, y) \in L_2 \quad (\text{in } x, y),$$

* That is, values δ_i ($i=0, \dots, n-1$) such that the italicized statement preceding (2.6b) holds.

so that in this case $K(x, y)$ will be a kernel for which the results of the Fredholm type will hold. *Kernels* $K(x, y) \in H_1$ are precisely Carleman's kernels of the type considered in (C; chap. 4). For any $n > 1$ it is possible to show that there exist kernels which belong to H_n and at the same time do not belong to H_{n-1} ; we shall give such an example for $n = 2$.

The following observations regarding kernels included in H_n are in order, it being understood that everywhere in the sequel the values δ_i ($i = 0, \dots, n-1$) are taken as "admissible" (cf. footnote to Definition 2.2).

The function

$$|K^{\delta_0, \delta_1, \dots, \delta_{n-1}}(x, y)|$$

is monotone non-decreasing as $\delta_{n-1} \rightarrow 0$; the limit

$$(2.12) \quad \lim_{\delta_{n-1}} K^{\delta_0, \delta_1, \dots, \delta_{n-1}}(x, y) = K^{\delta_0, \delta_1, \dots, \delta_{n-2}}(x, y)$$

exists and

$$(2.12a) \quad |K^{\delta_0, \dots, \delta_{n-1}}(x, y)| \leq |K^{\delta_0, \delta_1, \dots, \delta_{n-2}}(x, y)|.$$

In succession we obtain the limits

$$(2.13) \quad \lim_{\delta_{n-2}} K^{\delta_0, \dots, \delta_{n-2}}(x, y) = K^{\delta_0, \dots, \delta_{n-3}}(x, y), \dots, \\ \lim_{\delta_1} K^{\delta_0, \delta_1}(x, y) = K^{\delta_0}(x, y), \quad \lim_{\delta_0} K^{\delta_0}(x, y) = K(x, y).$$

It is also noted that

$$|K^{\delta_0, \dots, \delta_i}(x, y)| \leq |K^{\delta_0, \dots, \delta_{i-1}}(x, y)|,$$

and that the first member in this inequality is monotone non-decreasing as $\delta_i \rightarrow 0$; this can be asserted for $i = n-1, n-2, \dots, 0$. In view of (2.13) one may write

$$(2.14) \quad K(x, y) = \lim_{\delta_0} \lim_{\delta_1} \dots \lim_{\delta_{n-2}} \lim_{\delta_{n-1}} K^{\delta_0, \delta_1, \dots, \delta_{n-1}}(x, y),$$

where the order of the limiting processes, in general, cannot be interchanged. It is also observed that the functions of the second members of (2.12), (2.13) belong to the classes H_i as follows

$$(2.15) \quad K^{\delta_0, \delta_1, \dots, \delta_i}(x, y) \in H_{n-1-i} \quad (i = 0, 1, \dots, n-1).^*$$

The above considerations lead to the conclusion that *kernels* $K(x, y) \in H_n$ are also of rank n , according to the terminology of §1.

Example of $K(x, y) \in H_2$, but not belonging to H_1 (that is, not of Carleman's type). To construct such an example we shall take $a=0$, $b=1$ and define

* As indicated before, the class H_0 is identical with the class of functions L_2 (in two variables).

$K(x, y)$ by the relations

$$(2.16) \quad K(x, y) = g(x) \quad (\text{for } 0 \leq y < x),$$

$$(2.16a) \quad K(x, y) = g(y) \quad (\text{for } 0 \leq x < y),$$

the definition for $y=x$ being immaterial;

$$(2.16b) \quad g(x) = g_\nu(x) \quad (1/(\nu+1) < x < 1/\nu; \nu = 1, 2, \dots),$$

$$(2.16c) \quad g_\nu(x) = 0 \quad (1/(\nu+1) < x < \gamma_\nu = (2\nu+1)/(2\nu(\nu+1))),$$

$$g_\nu(x) = \frac{c_\nu^{1/2}}{(\nu^{-2} - x^2)^{1/2}} \quad (\gamma_\nu \leq x < 1/\nu; c_\nu > 0).$$

For this kernel the set E ((2.1)) consists of the points $0, 1/\nu$ ($\nu=1, 2, \dots$); the derived set E^1 ((2.1a)) will be $E^1 = (s_1)$ ($s_1=0$). Thus $E \in R_1$ (Definition 2.1). The set $\Delta^0(\delta_0)$ will consist of a single interval (cf. (2.2))

$$(2.17) \quad \Delta^0(\delta_0) = (0, \delta_0) \quad (0 < \delta_0 < 1),$$

where $\delta_0 \neq 1/i$ ($i=1, 2, \dots$). For some integer $m(\delta_0)$

$$(2.17a) \quad 1/(m(\delta_0) + 1) < \delta_0 < 1/m(\delta_0).$$

The set $\Delta^1(\delta_1)$ will consist of the intervals (cf. (2.3))

$$(2.17b) \quad \Delta^1(\delta_1) = (1/\nu - \delta_1, 1/\nu + \delta_1) \quad (\nu = 1, 2, \dots, m(\delta_0)),$$

where $0 < \delta_1 \leq \delta_1(\delta_0)$ with $\delta_1(\delta_0)$ denoting a positive number less than each of the two numbers

$$(2.17c) \quad 1/(2m(\delta_0)(m(\delta_0) - 1)), \quad 1/(m(\delta_0)) - \delta_0.$$

Then, by (2.11), (2.11a) and (2.11b), we have for $y < x$

$$(2.18) \quad K^{\delta_0, \delta_1}(x, y) = 0 \quad [x \text{ in } \Delta^0(\delta_0), \Delta^1(\delta_1) \ (\nu = 1, \dots, m(\delta_0)); \text{ cf. (2.17), (2.17b)}];$$

$$(2.18a) \quad K^{\delta_0, \delta_1}(x, y) = g(x) \quad [x \text{ in } (0, 1) - \Delta^0(\delta_0) - \sum \Delta^1(\delta_1)];$$

for $x > y$ the function of the first member of (2.18) is defined by symmetry. Whence by virtue of (2.16b), (2.16c), it is inferred that

$$(2.19) \quad \int_0^1 \int_0^1 |K^{\delta_0, \delta_1}(x, y)|^2 dx dy = 2 \int_{x=\delta_0}^{1/m(\delta_0)-\delta_1} \int_{y=0}^x g_{m(\delta_0)}^2(x) dy dx + \sum_{\nu=1}^{m(\delta_0)-1} 2 \int_{x=1/(\nu+1)+\delta_1}^{1/\nu-\delta_1} \int_{y=0}^x g_\nu^2(x) dy dx,$$

where

$$(2.19a) \quad 2 \int_{x=1/(v+1)+\delta_1}^{1/v-\delta_1} \int_{y=0}^x g^2(x) dy dx = -\lambda_v + \log T^{-1}(v, \delta_1) \\ [\lambda_v = -c_v \log(v^{-2} - \gamma^2); T(v, \delta_1) = (2\delta_1/v - \delta_1^2)^{c_v}].$$

Hence, for all admissible $\delta_0 (>0)$,

$$\iint |K^{\delta_0, \delta_1}(x, y)|^2 dx dy \rightarrow +\infty \quad (\text{as } \delta_1 \rightarrow 0),$$

and clearly $K^{\delta_0}(x, y)$ does not belong to L_2 (in x, y). Consequently it is clear that $K(x, y)$, as given by (2.16)–(2.16c), is a kernel satisfying the conditions of the italicized statement preceding (2.16). It is essential to note that for the example considered above the integral

$$\int_0^1 K^{\delta_0}(x, y) dy$$

diverges; in fact, convergence of this integral would have meant that $K(x, y)$ is essentially of Carleman's type.*

Some of the developments for integral equations whose kernels are included in H_n will be given with the aid of operators L specified as follows.

DEFINITION 2.3. Given a kernel $K(x, y) \in H_n$ (Definition 2.2), a linear operator $L_x(\xi | h(x))$ (ξ a parameter) will be said to be associated with $K(x, y)$ if

$$(2.20) \quad L_x(\xi | K(x, y)) \in L_2 \quad (\text{in } y);$$

$$(2.21) \quad |L_x(\xi | K^{\delta_0, \delta_1, \dots, \delta_{n-1}}(x, y))| < \gamma(\xi | y),$$

where $\gamma(\xi | y) \in L_2$ (in y) and $\gamma(\xi | y)$ is independent of $\delta_0, \delta_1, \dots, \delta_{n-1}$;

$$(2.22) \quad \lim_{\delta_{n-1}} L_x(\xi | K^{\delta_0, \dots, \delta_{n-1}}(x, y)) = L_x(\xi | K^{\delta_0, \dots, \delta_{n-2}}(x, y)), \\ \lim_{\delta_{n-2}} L_x(\xi | K^{\delta_0, \dots, \delta_{n-2}}(x, y)) = L_x(\xi | K^{\delta_0, \dots, \delta_{n-3}}(x, y)), \dots,$$

$$\lim_{\delta_0} L_x(\xi | K^{\delta_0}(x, y)) = L_x(\xi | K(x, y));$$

whenever $f_\nu(x) \in L_2$ converges weakly (as $\nu \rightarrow \infty$) to $f(x)$ ($a \leq x \leq b$) we have

$$(2.23) \quad \lim_{\nu} L_x(\xi | f_\nu(x)) = L_x(\xi | f(x));$$

$$(2.24) \quad \int_a^b L_x(\xi | K^{\delta_0, \dots, \delta_{n-1}}(x, y)) \phi(y) dy = L_x\left(\xi | \int_a^b K^{\delta_0, \dots, \delta_{n-1}}(x, y) \phi(y) dy\right),$$

whenever $\phi(x) \in L_2$.

* This follows by Carleman, Annales de l'Institut H. Poincaré, loc. cit.

NOTE. For $n=1$ an operator described in the above definition reduces precisely to the operator L given in (C, pp. 137, 138).

In order to make certain that those of the developments, with respect to kernels included in H_n ($n > 1$), which are made with the aid of operators L (Definition 2.3) should have a significance, it is essential to show the following.

There exist kernels $K(x, y)$, included in H_n ($n > 1$) and not belonging to H_{n-1} , with which one can associate an operator L satisfying the conditions of Definition 2.3.

We shall give such an example for $n=2$. For $n > 2$ similar examples can be given following similar procedures.* It will be sufficient to construct an operator L associated with the kernel $K(x, y)$, given by (2.16)–(2.16c). Let us take

$$(2.25) \quad L_x(\xi | h(x)) = \int_0^1 G(\xi | x) h(x) dx,$$

where, for $\nu = 1, 2, \dots$,

$$(2.25a) \quad G(\xi | x) = G_\nu(\xi | x) \quad (\text{for } \gamma_\nu \leq x < 1/\nu; \gamma_\nu \text{ from (2.16c)}),$$

$$(2.25b) \quad G(\xi | x) = -G_\nu(\xi | 2\gamma_\nu - x) \quad (\text{for } 1/(\nu+1) < x \leq \gamma_\nu);^\dagger$$

here we take

$$(2.25c) \quad G_\nu(\xi | x) = c_\nu^{-1/2}(\nu^{-2} - x^2)^{1/2} w_\nu(\xi | x), \quad 0 \leq w_\nu(\xi | x) \leq H,$$

where H is independent of ν , x and $w_\nu(\xi | x)$ [included in L_1 in x on $(\gamma_\nu, 1/\nu)$] is monotone non-increasing in x on $(\gamma_\nu, 1/\nu)$. Moreover, the c_ν will be taken subject to the requirement that the series

$$(2.26) \quad \sum_\nu \frac{1}{c_\nu \nu^5}$$

be convergent. We shall now demonstrate that the operator $L_x(\xi | h(x))$ ((2.25)), so defined, satisfies the conditions (2.20)–(2.24) with respect to the kernel $K(x, y)$ [(2.16)–(2.16c)].

By (2.25c)

$$|G_\nu(\xi | x)|^2 \leq \frac{H^2}{c_\nu} \left(\frac{1}{\nu^2} - \gamma_\nu^2 \right) < \frac{H^2}{c_\nu \nu^3} \quad (\gamma_\nu \leq x < 1/\nu);$$

thus, in view of (2.25a) and (2.25b),

$$(2.27) \quad |G(\xi | x)|^2 < \frac{H^2}{c_\nu \nu^3} \quad (1/(\nu+1) < x < 1/\nu; \nu = 1, 2, \dots).$$

* This will not be done in these pages in order to save space.

† γ_ν bisects the interval $(1/(\nu+1), 1/\nu)$; (2.25b) implies symmetry of $G(\xi | x)$ with respect to γ_ν , as indicated.

Hence

$$(2.27a) \quad \int_0^1 |G(\xi|x)|^2 dx = \sum_{r=1}^{\infty} \int_{1/(r+1)}^{1/r} |G(\xi|x)|^2 dx \\ < H^2 \sum_{r=1}^{\infty} \frac{1}{c_r r^3} \left(\frac{1}{r} - \frac{1}{r+1} \right) < H^2 \sum_r \frac{1}{c_r r^3};$$

the series last displayed being convergent in view of (2.26), it is concluded that

$$(2.28) \quad G(\xi|x) \in L_2 \quad (\text{in } x; 0 \leq x \leq 1).$$

By virtue of (2.16), (2.16a) and (2.25)

$$(2.29) \quad L_x(\xi|K(x,y)) = \beta(\xi|y) + \alpha(\xi|y),$$

where

$$(2.29a) \quad \beta(\xi|y) = g(y) \int_{x=0}^y G(\xi|x) dx, \quad \alpha(\xi|y) = \int_y^1 G(\xi|x) g(x) dx.$$

By (2.16b), (2.16c)

$$(2.30) \quad \beta(\xi|y) = 0 \quad (\text{for } 1/(r+1) < y < \gamma_r).$$

Now suppose

$$(2.31) \quad \gamma_r \leq y < \frac{1}{r};$$

then by (2.16c) from (2.29a) we deduce

$$\beta(\xi|y) = \frac{c_r^{1/2}}{(y^{-2} - y^2)^{1/2}} \left[\sum_{i=r+1}^{\infty} \int_{1/(i+1)}^{1/i} G(\xi|x) dx + \int_{1/(r+1)}^y G(\xi|x) dx \right];$$

in view of (2.25a) and (2.25b)

$$\int_{1/(i+1)}^{1/i} G(\xi|x) dx = 0 \quad (i = 1, 2, \dots);$$

whence

$$\begin{aligned} \beta(\xi|y) &= \frac{c_r^{1/2}}{(y^{-2} - y^2)^{1/2}} \int_{1/(r+1)}^y G(\xi|x) dx = \frac{c_r^{1/2}}{(y^{-2} - y^2)^{1/2}} \int_{1/(r+1)}^{2\gamma_r - y} G(\xi|x) dx \\ &= - \frac{c_r^{1/2}}{(y^{-2} - y^2)^{1/2}} \int_{1/(r+1)}^{2\gamma_r - y} G_r(\xi|2\gamma_r - x) dx \\ &= - \frac{c_r^{1/2}}{(y^{-2} - y^2)^{1/2}} \int_y^{1/r} G_r(\xi|u) du; \end{aligned}$$

in view of (2.25c) and in consequence of the monotone character of $w_\nu(\xi|u)$ it is concluded that, under (2.31), the integrand last displayed satisfies the inequality $0 \leq G_\nu(\xi|u) \leq G_\nu(\xi|y)$ ($y \leq u \leq 1/\nu$); thus, by (2.25c)

$$(2.32) \quad |\beta(\xi|y)| \leq \frac{c_\nu^{1/2}}{(\nu^2 - y^2)^{1/2}} G_\nu(\xi|y) \left(\frac{1}{\nu} - y \right) \leq \left(\frac{1}{\nu} - y \right) H < \frac{H}{2\nu^2} \quad (\text{under (2.31)}).$$

Inasmuch as (2.30), (2.32) hold for $\nu = 1, 2, \dots$, it is inferred that

$$(2.33) \quad |\beta(\xi|y)| < H/2 \quad (0 \leq y \leq 1).$$

On turning attention to $\alpha(\xi|y)$ ((2.29a)) it is found that

$$(2.34) \quad \begin{aligned} |\alpha(\xi|y)| &\leq \alpha(\xi) = \int_0^1 |G(\xi|x)| g(x) dx \\ &= \sum_{\nu=1}^{\infty} \int_{1/(\nu+1)}^{1/\nu} |G(\xi|x)| g_\nu(x) dx; \end{aligned}$$

by (2.16c), (2.25a) and (2.25c)

$$(2.34a) \quad \alpha(\xi) = \sum_{\nu=1}^{\infty} \int_{\gamma_\nu}^{1/\nu} |G_\nu(\xi|x)| g_\nu(\xi|x) dx = \sum_{\nu=1}^{\infty} \int_{\gamma_\nu}^{1/\nu} w_\nu(\xi|x) dx < H.$$

By virtue of (2.33), (2.34), (2.34a), on taking account of (2.29) it is deduced that (2.20) is satisfied for the example under consideration.

We shall now proceed to establish (2.21) (with $n=2$). By definition of $K^{\delta_0, \delta_1}(x, y)$ [(2.18), (2.18a)]

$$(2.35) \quad L_x(\xi|K^{\delta_0, \delta_1}(x, y)) = \beta^{\delta_0, \delta_1}(\xi|y) + \alpha^{\delta_0, \delta_1}(\xi|y),$$

where

$$(2.35a) \quad \beta^{\delta_0, \delta_1}(\xi|y) = g^{\delta_0, \delta_1}(y) \int_{x=0}^y G(\xi|x) dx,$$

$$(2.35b) \quad \begin{aligned} \alpha^{\delta_0, \delta_1}(\xi|y) &= \int_y^1 G(\xi|x) g^{\delta_0, \delta_1}(x) dx; \\ g^{\delta_0, \delta_1}(x) &= 0 \quad [\text{for } 0 \leq x \leq \delta_0; \text{ for } x \text{ on closed intervals} \\ &\quad (1/\nu - \delta_1, 1/\nu + \delta_1) (\nu = 1, 2, \dots, m(\delta_0))]; \\ g^{\delta_0, \delta_1}(x) &= g_\nu(x) \quad [1/(\nu+1) + \delta_1 < x < 1/\nu - \delta_1; \\ &\quad \nu = 1, 2, \dots, m(\delta_0) - 1; \text{ cf. (2.16c)}]; \\ g^{\delta_0, \delta_1}(x) &= g_{m(\delta_0)}(x) \quad [\delta_0 < x < 1/m(\delta_0) - \delta_1]. \end{aligned}$$

Clearly

$$(2.36) \quad 0 \leq g^{\delta_0, \delta_1}(x) \leq g^{\delta_0}(x) \leq g(x) \quad (0 \leq x \leq 1);$$

here

$$(2.36a) \quad g^{\delta_0}(x) = \lim_{\delta_1} g^{\delta_0, \delta_1}(x), \quad g(x) = \lim_{\delta_0} g^{\delta_0}(x).$$

By (2.35a) and (2.36) in view of (2.29a) and (2.33)

$$(2.37) \quad |\beta^{\delta_0, \delta_1}(\xi|y)| \leq g(y) \left| \int_{x=0}^y G(\xi|x) dx \right| = |\beta(\xi|y)| < \frac{H}{2}.$$

On the other hand, in consequence of (2.35a), (2.36), (2.34) and (2.34a),

$$(2.37a) \quad |\alpha^{\delta_0, \delta_1}(\xi|y)| \leq \int_y^1 |G(\xi|x)| g^{\delta_0, \delta_1}(x) dx \leq \int_y^1 |G(\xi|x)| g(x) dx < H.$$

In view of (2.35), (2.37) and (2.37a) it is inferred that *condition* (2.21) (Definition 2.3) holds with $\gamma(\xi|y) = 3H/2$.

To demonstrate the first one of the relations (2.22) it is sufficient to prove that

$$(2.38) \quad \lim_{\delta_1} \beta^{\delta_0, \delta_1}(\xi|y) = \beta^{\delta_0}(\xi|y), \quad \lim_{\delta_1} \alpha^{\delta_0, \delta_1}(\xi|y) = \alpha^{\delta_0}(\xi|y),$$

where

$$(2.38a) \quad \beta^{\delta_0}(\xi|y) = g^{\delta_0}(y) \int_{x=0}^y G(\xi|x) dx, \quad \alpha^{\delta_0}(\xi|y) = \int_y^1 G(\xi|x) g^{\delta_0}(x) dx,$$

with $g^{\delta_0}(y)$ denoting the first function displayed in (2.36a),

$$g^{\delta_0}(x) = 0 \quad (0 \leq x \leq \delta_0), \quad g^{\delta_0}(x) = g(x) \quad (\delta_0 < x \leq 1).$$

The first of the equalities (2.38) follows immediately from the first relations in (2.35a) and (2.36a). To justify the second relation in (2.38) it is sufficient to show that

$$\lim_{\delta_1} \int_y^1 G(\xi|x) g^{\delta_0, \delta_1}(x) dx = \int_y^1 G(\xi|x) g^{\delta_0}(x) dx.$$

The passage to the limit under the integral sign is here justified because the integrand displayed in the first member converges to the integrand displayed in the second member while, as follows by (2.36),

$$|G(\xi|x) g^{\delta_0, \delta_1}(x)| \leq |G(\xi|x)| g(x) \in L_1 \quad (\text{in } x; \text{ cf. (2.34), (2.34a)}).$$

The second one of the relations (2.22) will certainly hold if

$$(2.39) \quad \lim_{\delta_0} \beta^{\delta_0}(\xi|y) = \beta(\xi|y), \quad \lim_{\delta_0} \alpha^{\delta_0}(\xi|y) = \alpha(\xi|y),$$

where $\beta^{\delta_0}(\xi|y)$, $\alpha^{\delta_0}(\xi|y)$ are given by (2.38a) and $\beta(\xi|y)$, $\alpha(\xi|y)$ are the functions of (2.29a). The first of the equalities (2.39) is a consequence of the last one of (2.36a). The other equality of (2.39), that is the relation

$$\lim_{\delta_0} \int_y^1 G(\xi|x) g^{\delta_0}(x) dx = \int_y^1 G(\xi|x) g(x) dx,$$

is seen to be true in view of (2.36a) and of the inequality

$$|G(\xi|x) g^{\delta_0}(x)| \leq |G(\xi|x)| g(x) \in L_1 \quad (\text{in } x),$$

which is deduced from (2.36).

Accordingly it can be asserted that *conditions (2.22) of Definition 2.3 all hold* for the case under consideration.

The condition stated in connection with (2.23) *will hold* for all sequences $\{f_v(x)\}$ therein specified, since

$$\lim_v \int_0^1 G(\xi|x) f_v(x) dx = \int_0^1 G(\xi|x) f(x) dx;$$

in fact, passage to the limit under the integral sign is here justified in view of (2.28) and of Theorem 1.4.

It remains to verify whether (2.24) holds, that is whether we have

$$(2.40) \quad \int_{y=0}^1 \left[\int_{x=0}^1 G(\xi|x) K^{\delta_0, \delta_1}(x, y) dx \right] \phi(y) dy \\ = \int_{x=0}^1 G(\xi|x) \left[\int_{y=0}^1 K^{\delta_0, \delta_1}(x, y) \phi(y) dy \right] dx$$

for all $\phi(y) \in L_2$. The indicated change of order of integration can be justified without difficulty.

The developments from (2.25) to (2.40) enable us to conclude that the kernel $K(x, y)$, as given by (2.16)–(2.16c) and with the $c_v (>0)$ such that the series (2.26) converges, has associated with it an operator L (cf. (2.25)–(2.25c)) satisfying the conditions of Definition 2.3.

3. Formulation of induction for classes H_n . With $K(x, y) \in H_n$ (Definition 2.2) and $K^{\delta_0, \dots, \delta_{n-1}}(x, y)$ being the function specified by (2.11), (2.11a), (2.11b), consider equations

$$(3.1) \quad \phi^{\delta_0, \dots, \delta_{n-1}}(x) - \lambda \int_a^b K^{\delta_0, \dots, \delta_{n-1}}(x, y) \phi^{\delta_0, \dots, \delta_{n-1}}(y) dy = f(x) \quad (f(x) \in L_2),$$

$$(3.2) \quad \phi^{\delta_0, \dots, \delta_{n-1}}(x) - \lambda \int_a^b K^{\delta_0, \dots, \delta_{n-1}}(x, y) \phi^{\delta_0, \dots, \delta_{n-1}}(y) dy = 0.$$

By (2.10) the kernel in (3.1), (3.2) belongs to H_0 and is thus essentially a Fredholm kernel. In accordance with known facts regarding such equations, the spectrum of the kernel displayed in (3.2) is the function

$$(3.3) \quad \theta^{\delta_0, \dots, \delta_{n-1}}(x, y | \lambda) = \sum \phi_r^{\delta_0, \dots, \delta_{n-1}}(x) \phi_r^{\delta_0, \dots, \delta_{n-1}}(y) \\ (\lambda > 0; \text{summation over values } \nu \text{ such that } 0 < \lambda_{\nu}^{\delta_0, \dots, \delta_{n-1}} < \lambda);$$

$$(3.3a) \quad \theta^{\delta_0, \dots, \delta_{n-1}}(x, y | 0) = 0;$$

$$(3.3b) \quad \theta^{\delta_0, \dots, \delta_{n-1}}(x, y | \lambda) = - \sum \phi_r^{\delta_0, \dots, \delta_{n-1}}(x) \phi_r^{\delta_0, \dots, \delta_{n-1}}(y) \\ (\lambda < 0; \text{summation over values } \nu \text{ such that } \lambda \leq \lambda_{\nu}^{\delta_0, \dots, \delta_{n-1}} < 0).$$

Here the sequence

$$\{\phi_r^{\delta_0, \dots, \delta_{n-1}}(x)\}$$

forms an orthogonal normal set. The $\lambda_{\nu}^{\delta_0, \dots, \delta_{n-1}}$ are the characteristic values of (3.2); thus

$$(3.3c) \quad \phi_r^{\delta_0, \dots, \delta_{n-1}}(x) = \lambda_{\nu}^{\delta_0, \dots, \delta_{n-1}} \int_a^b K^{\delta_0, \dots, \delta_{n-1}}(x, y) \phi_r^{\delta_0, \dots, \delta_{n-1}}(y) dy.*$$

By induction we shall establish that certain facts, to be stated explicitly in the remainder of this section, hold for all integral equations (1.1), (1.2) whose kernels are included in H_m , where m is any finite integer (≥ 0). Thus, assume that the following facts, stated throughout the rest of this section, hold for kernels included in H_n ($n = 1, \dots, m-1$). An examination of these statements leads to the conclusion that they certainly hold true for $m=2$; that is, for Carleman kernels H_1 ; this can be asserted on the basis of (C; chap. 4). In subsequent sections these facts will be shown to hold for $n=m$; which will complete the induction.

Form the function

$$(3.4) \quad \int_a^x \int_a^y \theta^{\delta_0, \dots, \delta_{n-1}}(x, y | \lambda) dx dy = \Omega^{\delta_0, \dots, \delta_{n-1}}(x, y | \lambda) \text{ (cf. (3.3)-(3.3b))}.$$

Subsequences of positive numbers

$$(3.5) \quad \delta_{n-1,r} \ (r = 1, 2, \dots), \ \delta_{n-2,r} \ (r = 1, 2, \dots), \dots, \ \delta_{0,r} \ (r = 1, 2, \dots)$$

can be found so that

$$(3.5a) \quad \lim_r \delta_{n-1,r} = 0, \dots, \lim_r \delta_{0,r} = 0,$$

* Many properties of $\theta^{\delta_0, \dots, \delta_{n-1}}(x, y | \lambda)$ can be inferred from (C).

and so that the limits

$$\begin{aligned}
 (3.6) \quad & \lim_{\delta_{n-1}, r} \Omega^{\delta_0, \dots, \delta_{n-r}}(x, y | \lambda) = \Omega^{\delta_0, \dots, \delta_{n-2}}(x, y | \lambda), \\
 & \lim_{\delta_{n-2}, r} \Omega^{\delta_0, \dots, \delta_{n-2}}(x, y | \lambda) = \Omega^{\delta_0, \dots, \delta_{n-3}}(x, y | \lambda), \dots, \\
 & \lim_{\delta_0, r} \Omega^{\delta_0, r}(x, y | \lambda) = \Omega(x, y | \lambda)
 \end{aligned}$$

exist for all (x, y, λ) ,* convergence to the limits (3.6) being uniform with respect to (x, y) ;

$$(3.7) \quad \text{Var. } \Omega(x, y | \lambda) \leq [(x-a)(y-a)]^{1/2}; \dagger \quad \Omega(x, y | 0) = 0;$$

$$\begin{aligned}
 (3.7a) \quad & |\Omega(x', y' | \lambda) - \Omega(x, y | \lambda)| \\
 & \leq [(b-a)|y' - y|]^{1/2} + [(b-a)|x' - x|]^{1/2}.
 \end{aligned}$$

The function $\Omega(x, y | \lambda)$ may be discontinuous in λ for certain values of λ , say $\lambda_1, \lambda_2, \dots$.

We have

$$\begin{aligned}
 (3.8) \quad & \int_a^b \left| \frac{\partial}{\partial y} \Omega(x, y | \lambda) \right|^2 dy \leq x - a \\
 & \quad \quad \quad [\text{integrand exists for almost all } y; a \leq y \leq b].
 \end{aligned}$$

With the numbers (3.5) suitably chosen one has

$$\begin{aligned}
 (3.8a) \quad & \lim_r \frac{\partial}{\partial y} \Omega^{\delta_0, \dots, \delta_{n-1}, r}(x, y | \lambda) = \frac{\partial}{\partial y} \Omega^{\delta_0, \dots, \delta_{n-2}}(x, y | \lambda), \\
 & \lim_r \frac{\partial}{\partial y} \Omega^{\delta_0, \dots, \delta_{n-2}, r}(x, y | \lambda) = \frac{\partial}{\partial y} \Omega^{\delta_0, \dots, \delta_{n-3}}(x, y | \lambda), \dots, \\
 & \lim_r \frac{\partial}{\partial y} \Omega^{\delta_0, r}(x, y | \lambda) = \frac{\partial}{\partial y} \Omega(x, y | \lambda),
 \end{aligned}$$

convergence being in the *weak sense* in y ($a \leq y \leq b$).[‡] Also

$$(3.9) \quad \int_a^b \left| \frac{\partial}{\partial x} \int_{y=a}^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right|^2 dx \leq \int_a^b h^2(x) dx$$

* That is, "in general" for $a \leq x, y \leq b$ and for all real λ .

† "Var." means "variation with respect to λ " for $(-\infty, +\infty)$ unless the interval is indicated explicitly.

‡ (3.8) is assumed for kernels of classes H_1, H_2, \dots, H_{n-1} ; in particular (3.8) will hold for the second members of (3.8a), which are defined for almost all y on (a, b) .

whenever $h(x) \in L_2$; moreover,

$$(3.9a) \quad \lim_r \frac{\partial}{\partial x} \int_{y=a}^b h(y) \frac{\partial}{\partial y} \Omega^{\delta_{0,r}}(x, y | \lambda) dy = \frac{\partial}{\partial x} \int_{y=a}^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy,$$

convergence being *weak* in x (for a suitable sequence $\delta_{0,r}$).

Whenever $g(x), h(y) \in L_2$ the following relations will hold (provided the $\delta_{0,r}$ are suitably chosen):

$$(3.10) \quad \lim_r \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_{y=a}^b h(y) \frac{\partial}{\partial y} \Omega^{\delta_{0,r}}(x, y | \lambda) dy \right] dx \\ = \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right] dx;$$

$$(3.11) \quad \left| \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right] dx \right| \\ \leq \left[\int_a^b h^2(x) dx \right]^{1/2} \left[\int_a^b g^2(x) dx \right]^{1/2};$$

$$(3.11a) \quad \text{Var.} \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right] dx \\ \leq \text{second member above.}$$

Whenever $\alpha(\lambda)$ is continuous on (λ_1, λ_2) and $g(x), h(y) \in L_2$

$$(3.12) \quad \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega^{\delta_{0,r}}(x, y | \lambda) dy \right] dx \\ \rightarrow \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right] dx$$

(as $r \rightarrow 0$; suitable $\delta_{0,r}$).

With $\alpha(\lambda)$ continuous on (λ_1, λ_2) and $|\alpha(\lambda)| \leq M$,

$$(3.13) \quad \int_a^b \left[\frac{\partial}{\partial x} \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right]^2 dx \\ \leq M^2 \int_a^b h^2(x) dx = A.*$$

The following interchanges of limits are justifiable for kernels of classes H_1, H_2, \dots, H_{m-1} :

* We assume this for kernels of classes H_1, H_2, \dots, H_{m-1} ; for kernels H_1 this inequality follows by developments in (C), but is by no means obvious.

$$\begin{aligned}
 & \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right] dx \\
 (3.14) \quad &= \int_a^b g(x) \frac{\partial}{\partial x} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right] dx \\
 &= \int_a^b g(x) \frac{\partial}{\partial x} \left[\int_a^b \frac{\partial}{\partial y} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \Omega(x, y | \lambda) \right] h(y) dy \right] dx;
 \end{aligned}$$

for H_1 this is assured in (C, p. 135).

The generalized *Bessel's inequality* for kernels of classes H_n ($n < m$) is

$$(3.15) \quad \int_{-\infty}^{\infty} d\lambda \left[\int_a^b h(x) \left(\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right) dx \right] \leq \int_a^b h^2(x) dx$$

(whenever $h(x) \in L_2$).

Following the terminology of (C) one may call $\Omega(x, y | \lambda)$, corresponding to kernels H_n , *closed* in case (3.15) holds with the equality sign.

When $\Omega(x, y | \lambda)$ is closed then, for every $h(x) \in L_2$,

$$(3.16) \quad h(x) = \frac{d}{dx} \int_{-\infty}^{\infty} d\lambda \left[\int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right]$$

almost everywhere on (a, b) .

Suppose there is an operator L , as specified in Definition 2.3, associated with our kernel $K(x, y) \in H_n$. Consider the equations

$$(3.17) \quad L_x(\xi | \phi(x)) - \lambda \int_a^b L_x(\xi | K^{\delta_0, \dots, \delta_{n-1}}(x, y)) \phi(y) dy = L_x(\xi | f(x)),$$

$$(3.18) \quad L_x(\xi | \phi(x)) - \lambda \int_a^b L_x(\xi | K(x, y)) \phi(y) dy = L_x(\xi | f(x)),$$

derived on the basis of (3.1) and (1.1), respectively. The following holds.

With $\Gamma\lambda = \beta \neq 0$ and $\phi^{\delta_0, \dots, \delta_{n-1}}(x)$ denoting a solution of (3.1), the repeated limit, in the sense of weak convergence,

$$\begin{aligned}
 & \lim_{\delta_0, r} \lim_{\delta_1, r} \dots \lim_{\delta_{n-1}, r} \phi^{\delta_0, r, \delta_1, r, \dots, \delta_{n-1}, r}(x) = \phi(x) \quad (\text{suitable choice} \\
 (3.19) \quad & \text{of } \delta_{\nu, r} (> 0; r = 1, 2, \dots; \nu = 0, \dots, n-1); \lim_r \delta_{\nu, r} = 0)
 \end{aligned}$$

will exist and will constitute a solution of (3.18); moreover,

$$(3.19a) \quad \int_a^b |\phi^2(x)| dx \leq M = \frac{|\lambda|^2}{\beta^2} \int_a^b |f(x)|^2 dx.$$

Corresponding to every function $\Omega(x, y|\lambda)$ defined as in (3.6a) the equation (3.18) has a solution

$$(3.20) \quad \phi(x) = f(x) + \lambda \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_{\mu} \int_a^b f(y) \frac{\partial}{\partial y} \Omega(x, y|\mu) dy,$$

provided $\Gamma\lambda \neq 0$; this solution satisfies the inequality (3.19a).

Suppose $h(y) \in L_2$ and write (with $l > 0$)

$$(3.21) \quad \begin{aligned} \psi(x, l | \delta_0, \dots, \delta_{n-1}) \\ = \left(\int_l^{\infty} + \int_{-\infty}^{-l} \right) \frac{1}{\mu} d_{\mu} \int_a^b \delta^{\delta_0, \dots, \delta_{n-1}}(x, y|\mu) h(y) dy. \end{aligned}$$

With the $\delta_{i,r} [> 0; r = 1, 2, \dots; i = 0, \dots, n-1]$ suitably chosen,

$$(3.21a) \quad \begin{aligned} \psi(x, l | \delta_0, \dots, \delta_{n-1,r}) &\rightarrow \psi(x, l | \delta_0, \dots, \delta_{n-2}), \\ \psi(x, l | \delta_0, \dots, \delta_{n-2,r}) &\rightarrow \psi(x, l | \delta_0, \dots, \delta_{n-3}), \dots, \\ \psi(x, l | \delta_{0,r}) &\rightarrow \psi(x, l) \quad (\text{as } r \rightarrow \infty), \end{aligned}$$

convergence being in the weak sense in x ; moreover,

$$(3.21b) \quad \int_a^b \psi^2(x, l) dx \leq \frac{1}{l^2} \int_a^b h^2(x) dx,$$

$$(3.21c) \quad \lim_{l \rightarrow \infty} L_x(\xi | \psi(x, l)) = 0.$$

For kernels $K(x, y)$ of classes H_n ($n < m$) and $h(y) \in L_2$

$$(3.22) \quad \begin{aligned} \int_a^b L_x(\xi | K(x, y)) h(y) dy \\ = \int_{-l}^l d_{\mu} \int_a^b L_x(\xi | K(x, s)) \left[\frac{\partial}{\partial s} \int_a^b h(t) \frac{\partial}{\partial t} \Omega(s, t|\mu) dt \right] ds \\ + L_x(\xi | \psi(x, l)) \\ = \int_{-\infty}^{\infty} d_{\mu} \int_a^b L_x(\xi | K(x, s)) \left[\frac{\partial}{\partial s} \int_a^b h(t) \frac{\partial}{\partial t} \Omega(s, t|\mu) dt \right] ds. \end{aligned}$$

On writing (with $l > 0$)

$$(3.23) \quad w(x, l) = f(x) - \tau(x, l), \quad \tau(x, l) = \frac{\partial}{\partial x} \int_{-l}^l d_{\lambda} \int_a^b f(y) \frac{\partial}{\partial y} \Omega(x, y|\lambda) dy,$$

we have ((3.23a) being a consequence of (3.13))

$$(3.23a) \quad \int_a^b \tau^2(x, l) dx \leq \int_a^b f^2(y) dy = q,$$

$$(3.23b) \quad \int_a^b w^2(x, l) dx \leq 4q,$$

and

$$(3.23c) \quad w(x, l_1) \rightarrow w(x), \quad \int_a^b w^2(x) dx \leq 4q \quad (l_1 < l_2 < \dots),$$

convergence being in the *weak sense*; moreover, $w(y)$ satisfies the equation

$$(3.23d) \quad \int_a^b L_x(\xi | K(x, y)) w(y) dy = 0.$$

A consequence of the statements in connection with (3.23)–(3.23c) is the following. *If the equation*

$$(3.24) \quad \int_a^b L_x(\xi | K(x, y)) \phi(y) dy = 0 \quad (\phi(y) \in L_2)$$

has only the solution $\phi(\eta) = 0$ (almost everywhere), then every $f(x) \in L_2$ has the representation

$$(3.24a) \quad f(x) = \frac{d}{dx} \int_{-\infty}^{\infty} d\lambda \int_a^b f(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \quad (\text{almost everywhere}).$$

Let $\Delta q(\lambda) = q(\lambda') - q(\lambda'')$ (real $\lambda', \lambda'', \lambda' < \lambda''$); then for all kernels of classes H_n ($n < m$) and for all $h(y) \in L_2$

$$(3.25) \quad \begin{aligned} & L_x \left(\xi \left| \frac{\partial}{\partial x} \Delta \int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \lambda) dy \right. \right) \\ & - \int_a^b L_x(\xi | K(x, s)) \left[\frac{\partial}{\partial s} \int_{\lambda'}^{\lambda''} \mu d\mu \int_a^b h(y) \frac{\partial}{\partial y} \Omega(s, y | \mu) dy \right] ds = 0; \end{aligned}$$

in particular,

$$(3.25a) \quad \begin{aligned} & L_x \left(\xi \left| \frac{\partial}{\partial x} \Delta \Omega(x, y | \lambda) \right. \right) \\ & - \int_a^b L_x(\xi | K(x, s)) \left[\frac{\partial}{\partial s} \int_{\lambda'}^{\lambda''} \mu d\mu \Omega(s, y | \mu) \right] ds = 0. \end{aligned}$$

Also for $K(x, y) \in H_n$ ($n < m$)

$$(3.26) \quad \begin{aligned} & \int_a^b L_x(\xi | K(x, y)) h(y) dy \\ & = \int_{-\infty}^{\infty} \frac{1}{\mu} d\mu L_x \left(\xi \left| \frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega(x, y | \mu) dy \right. \right) \end{aligned}$$

for all $h(y) \in L_2$. Furthermore the following relation will hold, for $K(x, y) \in H_n$ ($n < m$),

$$(3.27) \quad \begin{aligned} & L_x \left(\xi \left| \frac{\partial}{\partial x} \int_{\lambda'}^{\lambda''} \mu d_\mu \Omega(x, y | \mu) \right. \right) \\ & - \lambda \int_a^b L_x(\xi | K(x, s)) \left[\frac{\partial}{\partial s} \int_{\lambda'}^{\lambda''} \mu d_\mu \Omega(s, y | \mu) \right] ds \\ & = L_x \left(\xi \left| \frac{\partial}{\partial x} \int_{\lambda'}^{\lambda''} (\mu - \lambda) d_\mu \Omega(x, y | \mu) \right. \right) \quad (\text{if } \Gamma\lambda = 0). \end{aligned}$$

4. Developments without the aid of operators L . In §3 we have assumed and have stated certain facts (refer to the text from (3.4) to (3.27)) for classes H_n ($n = 1, 2, \dots, m-1$); an examination of Carleman's work leads to the conclusion that *these statements certainly hold for Carleman's kernels H_1* . We shall now prove that the results asserted from (3.4) to (3.28) hold for kernels $K(x, y) \in H_m$, as well. This will establish the theory for kernels included in H_ν , where ν (> 0) is any finite integer.

Let

$$(4.1) \quad K_1(x, y) \in H_m,$$

the spectrum corresponding to $K_1^{\delta_0, \dots, \delta_{m-1}}(x, y)$ being the function defined in (3.3), (3.3a), (3.3b), with $n = m$. In the definition of the spectrum are involved numbers $\lambda^{\delta_0, \dots, \delta_{m-1}}$ and functions $\phi^{\delta_0, \dots, \delta_{m-1}}(x)$ (orthogonal normal set), satisfying equation (3.3c), where we now write

$$n = m, \quad K^{\delta_0, \dots, \delta_{n-1}}(x, y) = K_1^{\delta_0, \dots, \delta_{m-1}}(x, y).$$

If one forms the function

$$(4.2) \quad \Omega_1^{\delta_0, \dots, \delta_{m-1}}(x, y | \lambda) = \int_a^x \int_a^y \theta_1^{\delta_0, \dots, \delta_{m-1}}(x, y | \lambda) dx dy$$

(cf. (3.4)) and notes that this function is a $\Omega(x, y | \lambda)$ belonging to H_0 , it is observed that in consequence of (3.5)–(3.6), there exist limits

$$\lim_{\delta_{m-1}, r} \Omega_1^{\delta_0, \dots, \delta_{m-1}, r}(x, y | \lambda) = \Omega_1^{\delta_0, \dots, \delta_{m-2}}(x, y | \lambda),$$

$$\lim_{\delta_{m-2}, r} \Omega_1^{\delta_0, \dots, \delta_{m-2}, r}(x, y | \lambda) = \Omega_1^{\delta_0, \dots, \delta_{m-3}}(x, y | \lambda), \dots,$$

$$\lim_{\delta_1, r} \Omega_1^{\delta_0, \delta_1, r}(x, y | \lambda) = \Omega_1^{\delta_0}(x, y | \lambda)$$

$$[\lim_r \delta_{i,r} = 0; i = m-1, m-2, \dots, 1].$$

The latter limit is a $\Omega(x, y|\lambda)$ -function belonging to the class H_{m-1} . This function, accordingly, satisfies (3.7), (3.7a). Whence the "Compactness Theorem" (§2) can be applied, thus enabling one to assert that

$$(4.3) \quad \lim_{\delta_{0,r}} \Omega_1^{\delta_{0,r}}(x, y|\lambda) = \Omega_1(x, y|\lambda) \quad (\text{suitable } \delta_{0,r}; r = 1, 2, \dots)$$

exists, with the limiting function satisfying (3.7), (3.7a). We note that $\Omega_1(x, y|\lambda)$ is a $\Omega(x, y|\lambda)$ -function belonging to our kernel (4.1) and that it may be discontinuous in λ for, say, $\lambda'_1, \lambda'_2, \dots$.

We supposed that (3.8) holds for Ω -functions belonging to H_{m-1} ; thus

$$(4.4) \quad \int_a^b \left| \frac{\partial}{\partial y} \Omega_1^{\delta_{0,r}}(x, y|\lambda) \right|^2 dy \leq x - a;$$

by Theorem 1.3 and in view of (4.3)

$$\frac{\partial}{\partial y} \Omega_1^{\delta_{0,r}}(x, y|\lambda) \rightarrow \frac{\partial}{\partial y} \Omega_1(x, y|\lambda) \quad (\text{as } r \rightarrow \infty; \text{suitable } \delta_{0,r} > 0),$$

convergence being in the weak sense in y ; moreover,

$$(4.4a) \quad \int_a^b \left| \frac{\partial}{\partial y} \Omega_1(x, y|\lambda) \right|^2 dy \leq x - a.$$

These considerations enable us to assert (3.8), (3.8a) for the class H_m .

By (3.9), stated for $\Omega_1^{\delta_{0,r}}(x, y|\lambda)$, with the aid of Theorem 1.3 we obtain the relation

$$(4.5) \quad \lim_r \frac{\partial}{\partial x} \int_{y=a}^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_{0,r}}(x, y|\lambda) dy = \frac{\partial}{\partial x} \int_{y=a}^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y|\lambda) dy$$

[weak convergence in x ; suitable $\delta_{0,r}$ ($r = 1, 2, \dots$)];

moreover, (3.9) will hold for $\Omega_1(x, y|\lambda)$.

With the aid of Theorem 1.4, in view of (4.5) and on writing

$$f_r(x) = \frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_{0,r}}(x, y|\lambda) dy,$$

we get (whenever $g(x), h(x) \in L_2$)

$$(4.6) \quad \lim_r \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_{0,r}}(x, y|\lambda) dy \right] dx$$

$$= \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y|\lambda) dy \right] dx,$$

which is (3.10) for $\Omega(x, y|\lambda)$.

Formula (3.11a) will hold in particular for $\Omega_1^{\delta_0, r}(x, y|\lambda)$; on taking account of (4.6) it is concluded that (3.11a) holds in the limit, that is with $\Omega(x, y|\lambda)$ replaced by $\Omega_1(x, y|\lambda)$. The inequality thus obtained enables us to assert that (3.11) will hold also for $\Omega_1(x, y|\lambda)$.

By (4.6), (3.11a), with $\Omega = \Omega_1^{\delta_0, r}$, in virtue of Theorem 1.1 it is deduced that

$$(4.7) \quad \lim_r \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y|\lambda) dy \right] dx \\ = \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y|\lambda) dy \right] dx$$

(whenever $\alpha(\lambda)$ is continuous and $g(x), h(y) \in L_2$); that is, (3.12) holds for $\Omega_1(x, y|\lambda)$.

In (4.7) replace x by t and let

$$(4.8) \quad g(t) = 1 \quad (a \leq t \leq x), \quad g(t) = 0 \quad (x < t \leq b);$$

then it is deduced that

$$(4.9) \quad \lim_r \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y|\lambda) dy \\ = \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y|\lambda) dy \\ (\alpha(\lambda) \text{ continuous, } h(y) \in L_2);$$

the relationship (4.9) will hold also for Ω -functions belonging to classes H_r ($r < m$). Now, we may write (3.13) for $\Omega_1^{\delta_0, r}(x, y|\lambda)$; the inequality so obtained, together with Theorem 1.3, would imply that, if the δ_0, r are suitably chosen,

$$(4.9a) \quad \frac{\partial}{\partial x} \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y|\lambda) dy \rightarrow \Gamma(x, \lambda) \quad (\text{as } r \rightarrow \infty),$$

convergence being in the weak sense (in x); by (4.9)

$$(4.9b) \quad \Gamma(x, \lambda) = \frac{\partial}{\partial x} \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y|\lambda) dy,$$

and (cf. (3.13)), in accordance with Theorem 1.3,

$$\int |\Gamma(x, \lambda)|^2 dx \leq A.$$

Thus it is concluded that (3.13) holds for the class H_m ; it is also clear that (4.9a), (4.9b) hold for all classes H_r ($r \leq m$).

We now proceed to establish the first equality (3.14) for Ω_1 . With $\Omega_1^{\delta_0, r}$ belonging to H_{m-1} , this equality takes the form

$$(4.10) \quad \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y | \lambda) dy \right] dx \\ = \int_a^b g(x) \frac{\partial}{\partial x} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y | \lambda) dy \right] dx.$$

In view of (3.13) (for $\Omega_1^{\delta_0, r}$), (4.9a), and (4.9b), application of Theorem 1.4 will yield the result

$$(4.10a) \quad \lim_r \int_a^b g(x) \frac{\partial}{\partial x} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y | \lambda) dy \right] dx \\ = \int_a^b g(x) \frac{\partial}{\partial x} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \right] dx \\ \text{[whenever } g(x) \in L_2; \text{ suitable } \delta_{0, r} (r = 1, 2, \dots)].$$

We shall have

$$(4.10b) \quad \lim_r \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y | \lambda) dy \right] dx \\ = \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \right] dx,$$

if it is shown that

$$(1) \quad \lim_r \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y | \lambda) dy \right] dx \\ = \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \right] dx,$$

and that

$$(2) \quad \text{Var.} \int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y | \lambda) dy \right] dx \leq B,$$

where B is independent of $\delta_{0, r}$; this it is possible to assert in consequence of Theorem 1.1. Now, (1) holds in view of (3.10) (with $\Omega = \Omega_1$); on the other hand, (2) is implied by (3.11a) (inasmuch as $\Omega_1^{\delta_0, r}$ is of class H_{m-1}). Thus (4.10b) is seen to be true; together with (4.10a) this relation enables us to deduce from (4.10) that the first equality of (3.14) holds for $\Omega = \Omega_1$.

In view of (4.10a) the second equality (3.14) will be established for $\Omega = \Omega_1$, provided it is shown that

$$\begin{aligned}
 (3) \quad \lim_r \int_a^b g(x) \frac{\partial}{\partial x} \left[\int_a^b \frac{\partial}{\partial y} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \Omega_1^{\delta_0, r}(x, y | \lambda) \right] h(y) dy \right] dx \\
 = \int_a^b g(x) \frac{\partial}{\partial x} \left[\int_a^b \frac{\partial}{\partial y} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \Omega_1(x, y | \lambda) \right] h(y) dy \right] dx.
 \end{aligned}$$

If one equates the first and the last member of (3.14), writing

$$\Omega = \Omega_1^{\delta_0, r}, \quad g = 1 \quad (\text{on } (a, x)), \quad g = 0 \quad (\text{on } (x, b)),$$

it is deduced that

$$\begin{aligned}
 (4.10c) \quad \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y | \lambda) dy \\
 = \int_a^b \frac{\partial}{\partial y} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \Omega_1^{\delta_0, r}(x, y | \lambda) \right] h(y) dy;
 \end{aligned}$$

by (3.13) (with $\Omega = \Omega^{\delta_0, r}$) the latter equality implies that

$$(4.10d) \quad \int_a^b \left[\frac{\partial}{\partial x} \int_a^b \frac{\partial}{\partial y} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \Omega_1^{\delta_0, r}(x, y | \lambda) \right] h(y) dy \right]^2 dx \leq A.$$

In view of (4.10d) and in consequence of Theorems 1.4, 1.3 it is observed that (3) and hence *the second equality* (3.14) will hold for Ω_1 , provided that

$$\begin{aligned}
 (4) \quad \lim_r \int_a^b \frac{\partial}{\partial y} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \Omega_1^{\delta_0, r}(x, y | \lambda) \right] h(y) dy \\
 = \int_a^b \frac{\partial}{\partial y} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \Omega_1(x, y | \lambda) \right] h(y) dy \quad (\text{suitable } \delta_{0, r}).
 \end{aligned}$$

In virtue of (4.10c) it is concluded that (4) will hold if

$$\begin{aligned}
 \lim_r \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y | \lambda) dy \\
 = \int_a^b \frac{\partial}{\partial y} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \Omega_1(x, y | \lambda) \right] h(y) dy;
 \end{aligned}$$

that is, in view of (4.9), if

$$\begin{aligned}
 (4.10e) \quad \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \\
 = \int_a^b \frac{\partial}{\partial y} \left[\int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d_\lambda \Omega_1(x, y | \lambda) \right] h(y) dy.
 \end{aligned}$$

Now, (4.10e) can be established with the aid of (4.10c). In fact, by (4.9) the first member of (4.10c) will yield in the limit the first member of (4.10e); on the other hand, for suitable $\delta_{0,r}$ ($r=1, 2, \dots$)

$$(4.11) \quad \lim_r \int_a^b h(y) \left[\frac{\partial}{\partial y} \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \Omega_1^{\delta_{0,r}}(x, y | \lambda) \right] dy = \text{second member of} \quad (4.10e),$$

because, by (3.13) (with $\Omega = \Omega_1^{\delta_{0,r}}$ and $h=1$ on (a, y) and $h=0$ on (y, b)),

$$(4.11a) \quad \int_a^b \left[\frac{\partial}{\partial x} \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \Omega_1^{\delta_{0,r}}(x, y | \lambda) \right]^2 dx \leq A \quad (r=1, 2, \dots),^*$$

and since

$$(4.11b) \quad \lim_r \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \Omega_1^{\delta_{0,r}}(x, y | \lambda) = \int_{\lambda_1}^{\lambda_2} \alpha(\lambda) d\lambda \Omega_1(x, y | \lambda).^\dagger$$

To ascertain the truth of (4.11) on the basis of (4.11a) and (4.11b) one needs only to take note of Theorem 1.3 and of Theorem 1.4. With (4.10e) established we have (4) secured, as well as the *second equality* (3.14) (for Ω_1).

Thus, (3.14) holds for the class H_m .

To establish the generalized Bessel's inequality for Ω_1 it is observed first that, in consequence of (3.11a) with

$$\Omega = \Omega_1^{\delta_{0,r}}, \quad g(x) = h(x) \in L_2,$$

it follows that

$$(4.12) \quad \text{Var.} \int_a^b h(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_{0,r}}(x, y | \lambda) dy \right] dx \leq \int_a^b h^2(x) dx.$$

Also

$$(4.12a) \quad \lim_r \int_a^b h(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_{0,r}}(x, y | \lambda) dy \right] dx \\ = \int_a^b h(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \right] dx,$$

which is deduced from (3.10) (for $\Omega = \Omega_1$ and $g(x) = h(x)$). In view of (4.12) and (4.12a), with the aid of Theorem 1.1, and on writing (3.15) (with $\Omega = \Omega_1^{\delta_{0,r}}$), and on letting $r \rightarrow \infty$, it is inferred without difficulty that

* Here x and y may be interchanged.

† This relation is a consequence of the inequality $\text{Var.} \Omega_1^{\delta_{0,r}} \leq [(x-a)(y-a)]^{1/2}$ and of the Theorem 1.1; (4.11b) also follows (4.9).

$$(4.13) \quad \int_{-\infty}^{\infty} d_{\lambda} \left[\int_a^b h(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \right] dx \right] \leq \int_a^b h^2(x) dx,$$

which is the desired inequality.

When Ω_1 is *closed*, so that (4.13) holds with the equality sign, we obtain the representation (3.16), with $\Omega = \Omega_1$, by a device of the type employed in (C). That is, replace $h(x)$ in (4.13) by $h(x) + g(x)$, obtaining

$$(4.14) \quad \begin{aligned} 2 \int_a^b h(x)g(x)dx = & \int_{-\infty}^{\infty} d_{\lambda} \left[\int_a^b h(x) \left[\frac{\partial}{\partial x} \int_a^b g(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \right] dx \right] \\ & + \int_{-\infty}^{\infty} d_{\lambda} \left[\int_a^b g(x) \left[\frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \right] dx \right]. \end{aligned}$$

In the first term of the second member of (4.14) interchange x and y and then let

$$g = 1 \quad (\text{on } (a, x)), \quad g = 0 \quad (\text{on } (x, b));$$

the representation (3.16) (with $\Omega = \Omega_1$) will result immediately.

Thus, all the statements which have been made in §3 up to (3.16), inclusive, hold for the class H_m , as well. The statements of §3, just referred to, have been made for classes H_n ($n < m$) without the use of operators L (Definition 2.3). The results therein indicated have been extended in the above to the class H_m ; in the process of the extension operators L have not been employed. Hence the induction is complete with respect to the statements, in question, of §3. We state this result as follows.

THEOREM 4.1. *With classes H_n specified by Definition 2.2, all the statements made in §3 up to (3.16) (inclusive) will hold true for all classes H_n ($n = 1, 2, \dots$).*

5. Developments on the basis of operators L . Let L' be an operator as specified in Definition 2.3 and supposed to exist, associated with our kernel $K_1(x, y) \in H_m$. We form the equation

$$(5.1) \quad L'_z(\xi | \phi_1(x)) - \lambda \int_a^b L'_z(\xi | K_1(x, y)) \phi_1(y) dy = L'_z(\xi | f(x))$$

(with given $f(x) \in L_2$), as well as the related equation

$$(5.2) \quad L'_z(\xi | \phi_1^{\delta_0, r}(x)) - \lambda \int_a^b L'_z(\xi | K_1^{\delta_0, r}(x, y)) \phi_1^{\delta_0, r}(y) dy = L'_z(\xi | f(x)).$$

It is essential to demonstrate that the operator L' is "associated" (in the sense of Definition 2.3) with the kernel $K_1^{\delta_0, r}(x, y) \in H_{m-1}$. Thus the following relations are to be verified, with $K^1 = K_1^{\delta_0}$:

$$(1) \quad L'_x(\xi | K^1(x, y)) \subset L_2 \quad (\text{in } y);$$

$$(2) \quad |L'_x(\xi | K^{1\delta_1, \dots, \delta_{m-1}}(x, y))| < \gamma(\xi | y) \subset L_2 \quad (\text{in } y);$$

$$L'_x(\xi | K^{1\delta_1, \dots, \delta_{m-1}}(x, y)) \xrightarrow{(\delta_{m-1})} L'_x(\xi | K^{1\delta_1, \dots, \delta_{m-2}}(x, y))$$

$$(3) \quad \xrightarrow{(\delta_{m-2})} L'_x(\xi | K^{1\delta_1, \dots, \delta_{m-3}}(x, y))$$

$$\xrightarrow{(\delta_1)} \dots L'_x(\xi | K^{1\delta_1}) \xrightarrow{(\delta_1)} L'_x(\xi | K^1(x, y));$$

$$(4) \quad L'_x(\xi | f_\nu(x)) \xrightarrow{(\nu)} L'_x(\xi | f(x)), \quad \text{if } f_\nu(x) \rightarrow f(x) \text{ in the weak sense;}$$

$$(5) \quad \int_a^b L'_x(\xi | K^{1\delta_1, \dots, \delta_{m-1}}(x, y)) \phi(y) dy = L'_x\left(\xi \left| \int_a^b K^{1\delta_1, \dots, \delta_{m-1}}(x, y) \phi(y) dy \right.\right)$$

(whenever $\phi(y) \subset L_2$).

If we designate by (2.20')–(2.24') the conditions (2.20)–(2.24), with K , L and n replaced by K_1 , L' , and m , the truth of (1)–(5) is inferred as follows. Conditions (2), (4), (5) are precisely the conditions (2.21'), (2.23'), (2.24'). The relations (3) are identical with those of (2.22'), the last limiting relation in (2.22') being omitted. As to (1), it is observed that, in view of (2) and (3),

$$|L'_x(\xi | K^1(x, y))| \leq \gamma(\xi | y) \subset L_2 \quad (\text{in } y),$$

which, together with other considerations, establishes (1).

By (3.19) and (3.19a) the equation (5.2) has a solution $^*\phi_{1^{\delta_0, r}}(x)$, such that

$$(5.3) \quad \int_a^b |\phi_{1^{\delta_0, r}}(x)|^2 dx \leq \frac{|\lambda|^2}{\beta^2} \int_a^b |f(x)|^2 dx = M \quad (\text{if } \Gamma\lambda = \beta \neq 0).$$

In consequence of Theorem 1.3, applicable in view of (5.3), we have for suitable $\delta_{0, r} (> 0; r = 1, 2, \dots; \lim_r \delta_{0, r} = 0)$

$$(5.4) \quad \lim_r \phi_{1^{\delta_0, r}}(x) = \phi_1(x),$$

convergence being in the *weak sense*; moreover,

$$(5.4a) \quad \int_a^b |\phi_1(x)|^2 dx \leq M \quad (M \text{ from (5.3)}).$$

It remains to demonstrate that $\phi_1(x)$ is a solution of (5.1). Substitute the function $\phi_{1^{\delta_0, r}}(x)$ (referred to in (5.3)) in (5.2) and let $r \rightarrow \infty$. We shall have

$$(5.5) \quad \lim_r L'_x(\xi | \phi_{1^{\delta_0, r}}(x)) = L'_x(\xi | \phi_1(x))$$

* This solution is obtained as a repeated limit, according to (3.19).

by (5.4) and (2.23'). On the other hand,

$$(5.5a) \quad \lim_r \int_a^b L_z'(\xi | K_1^{b_0, r}(x, y)) \phi_1^{b_0, r}(y) dy = \int_a^b L_z'(\xi | K_1(x, y)) \phi_1(y) dy.$$

This is a consequence of (5.3), (5.4), of the last limiting relation (2.22'), and of the inequality

$$|L_z'(\xi | K_1^{b_0, r}(x, y))| \leq \gamma(\xi | y) \in L_2 \quad (\text{in } y);^*$$

in fact, these conditions enable application of Theorem 1.5. In virtue of (5.5) and (5.5a) it can be asserted that the function $\phi_1(x)$, defined in (5.4) constitutes a solution of (5.1) (for $\beta \neq 0$). In view of the definition of $\phi_1^{b_0, r}(x)$ and in consequence of (5.4) it is observed that $\phi_1(x)$ is a repeated limit. The statement in connection with (3.19), (3.19a) can thus be made for the class H_m .

The important formula (3.20) will be extended to our equation (5.1) with the aid of the following relation

$$(5.6) \quad \begin{aligned} \lim_r \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_\mu \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1^{b_0, r}(x, y | \mu) dy \\ = \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_\mu \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1(x, y | \mu) dy, \end{aligned}$$

which we shall now proceed to prove. Let us write (with $l > |\text{real part of } \lambda|$)

$$(5.7) \quad \begin{aligned} \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_\mu p_r(x, \mu) &= \int_{-l}^l \frac{1}{\mu - \lambda} d_\mu p_r(x, \mu) + R_{l, r}(x, \lambda), \\ R_{l, r}(x, \lambda) &= \left(\int_l^{\infty} + \int_{-\infty}^{-l} \right) \frac{1}{\mu - \lambda} d_\mu p_r(x, \mu), \\ p_r(x, \mu) &= \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1^{b_0, r}(x, y | \mu) dy, \\ R_l(x, \lambda) &= \left(\int_l^{\infty} + \int_{-\infty}^{-l} \right) \frac{1}{\mu - \lambda} d_\mu \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1(x, y | \mu) dy. \end{aligned}$$

By (3.11a), with

$$\Omega = \Omega_1^{b_0, r}, \quad h(y) = f(y), \quad g = 1 \text{ (on } (a, x)), \quad g = 0 \text{ (on } (x, b)),$$

we have

$$(5.7a) \quad \text{Var. } p_r(x, \mu) \leq \left[\int_a^b f^2(x) dx \right]^{1/2} (x - a)^{1/2} \leq A.$$

* This is the relation obtained immediately preceding (5.3).

On the other hand, because of (3.9a) (with $\Omega = \Omega_1$)

$$(5.7b) \quad \lim_r p_r(x, \mu) = \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1(x, y | \mu) dy = p(x, \mu).$$

In view of (5.7a)

$$(5.7c) \quad \text{Var. } p(x, \mu) \leq A.$$

In consequence of (5.7a) and (5.7b), application of Theorem 1.1 will yield the result

$$(5.8) \quad \lim_r \int_{-l}^l \frac{1}{\mu - \lambda} d_\mu p_r(x, \mu) = \int_{-l}^l \frac{1}{\mu - \lambda} d_\mu p(x, \mu).$$

It is also noted that, by (5.7), (5.7a) and (5.7c),

$$|R_{l,r}(x, \lambda)| \quad \text{and} \quad |R_l(x, \lambda)| < \left(\frac{1}{|l + \lambda|} + \frac{1}{|l - \lambda|} \right) A.$$

Thus, for $\epsilon (>0)$ however small,

$$(5.8a) \quad |R_{l,r}(x, \lambda)| < \epsilon/3, \quad |R_l(x, \lambda)| < \epsilon/3,$$

provided l_* is taken sufficiently great. We have, for x and λ fixed ($\beta \neq 0$),

$$(5.9) \quad \left| \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_\mu p(x, \mu) - \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_\mu p_r(x, \mu) \right| = |R_l(x, \lambda) - R_{l,r}(x, \lambda)| \\ + \left[\int_{-l}^l \frac{1}{\mu - \lambda} d_\mu p(x, \mu) - \int_{-l}^l \frac{1}{\mu - \lambda} d_\mu p_r(x, \mu) \right] < \epsilon,$$

provided $l = l_*$ is such that (5.8a) holds and provided $r = r_*(x, \lambda)$ is taken sufficiently great (cf. (5.8)). *This establishes (5.6).*

We come now to the consideration of (3.20). In consequence of (3.20), applied to $\Omega_1^{\beta_0, r}$, it is concluded that a solution of (5.2) may be given in the form

$$(5.10) \quad \phi_1^{\beta_0, r}(x) = f(x) + \lambda \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_\mu \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1^{\beta_0, r}(x, y | \mu) dy \\ \text{(for } \beta \neq 0),$$

with

$$(5.10a) \quad \int_a^b |\phi_1^{\beta_0, r}(x)|^2 dx \leq M.$$

In virtue of (5.10a) and with the aid of a reasoning of the type previously employed in connection with (5.3)–(5.5a) it is concluded that

$$(5.11) \quad \lim_r \phi_1^{\delta_0, r}(x) = \phi_1(x), \quad \int_a^b |\phi_1(x)|^2 dx \leq M$$

(suitable δ_0, r ($r=1, 2, \dots$); convergence in the weak sense), where $\phi_1(x)$ is a solution of (5.1). Now (5.11) implies that (cf. (5.10))

$$\int_a^x \phi_1^{\delta_0, r}(x) dx = \int_a^x f(x) dx + \lambda \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_\mu \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y | \mu) dy$$

$\rightarrow H(x)$ [as $r \rightarrow \infty$; $H(x)$ absolutely continuous],

where

$$(5.11a) \quad \frac{d}{dx} H(x) = \phi_1(x) \quad [\phi_1(x) \text{ from (5.11); almost everywhere}].$$

Clearly, because of (5.6),

$$(5.11b) \quad H(x) = \int_a^x f(x) dx + \lambda \int_{-\infty}^{\infty} \frac{1}{\mu - \lambda} d_\mu \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1(x, y | \mu) dy.$$

From (5.11a) and (5.11b) it is deduced that $\phi_1(x)$ is represented by the formula (3.20) (with $\Omega = \Omega_1$). On taking account of (5.11) it is finally concluded that the italicized statement made in connection with (3.20) holds for the class H_m .

In accordance with (3.21) we write

$$(5.12) \quad \begin{aligned} &\psi_1(x, l | \delta_0, \dots, \delta_{m-1}) \\ &= \left(\int_l^\infty + \int_{-\infty}^{-l} \right) \frac{1}{\mu} d_\mu \int_a^b \theta_1^{\delta_0, \dots, \delta_{m-1}}(x, y | \mu) h(y) dy \\ &\quad [h(y) \in L_2; \theta_1^{\delta_0, \dots, \delta_{m-1}} = \text{spectrum of } K_1^{\delta_0, \dots, \delta_{m-1}}(x, y)]. \end{aligned}$$

By (3.21a), applied to ψ_1 of (5.12), one may assert only the following:

$$(5.12a) \quad \begin{aligned} &\psi_1(x, l | \delta_0, \dots, \delta_{m-1, r}) \rightarrow \psi_1(x, l | \delta_0, \dots, \delta_{m-2}), \\ &\psi_1(x, l | \delta_0, \dots, \delta_{m-2, r}) \rightarrow \psi_1(x, l | \delta_0, \dots, \delta_{m-3}), \dots, \\ &\psi_1(x, l | \delta_0, \delta_{1, r}) \rightarrow \psi_1(x, l | \delta_0), \end{aligned}$$

the $\delta_{r, r}$ ($r=m-1, \dots, 1$; $r=1, 2, \dots$) being suitably chosen and convergence being in the *weak sense* in x (as $r \rightarrow \infty$); moreover, in consequence of (3.21b) and (3.21c)

$$(5.12b) \quad \int_a^b \psi_1^2(x, l | \delta_0) dx \leq \frac{1}{l^2} \int_a^b h^2(x) dx,$$

$$(5.12c) \quad \lim_l L_x'(\xi | \psi_1(x, l | \delta_0, r)) = 0.$$

In virtue of (5.12b) with the aid of Theorem 1.3 it is deduced that

$$(5.13) \quad \lim_r \psi_1(x, l | \delta_{0,r}) = \psi_1(x, l)$$

[convergence in the weak sense; suitable $\delta_{0,r} (>0) \rightarrow 0$],

$$(5.13a) \quad \int_a^b \psi_1^2(x, l) dx \leq \frac{1}{l^2} \int_a^b h^2(x) dx.$$

From (5.13a) it is inferred that

$$\left| \int_a^x \psi_1(x, l) dx \right| \leq (b-a)^{1/2} \left[\int_a^x \psi_1^2(x, l) dx \right]^{1/2} \leq \frac{1}{l} A,$$

so that

$$\lim_l \int_a^x \psi_1(x, l) dx = 0.$$

Thus, $\psi_1(x, l)$ converges weakly (in x) to zero, as $l \rightarrow \infty$. Hence, in view of property (2.23')

$$(5.13b) \quad \lim_l L'(\xi | \psi_1(x, l)) = L'(\xi | 0) = 0.$$

The relations (5.13), (5.13a), (5.13b) imply that the statements made in connection with (3.21)–(3.21c) hold true for the class H_m .

By (3.22), applied to the kernel $K_1^{\delta_{0,r}}(x, y)$,

$$(5.14) \quad \begin{aligned} & \int_a^b L'_x(\xi | K_1^{\delta_{0,r}}(x, y)) h(y) dy \\ &= \int_{-l}^l d_\mu \int_a^b L'_x(\xi | K_1^{\delta_{0,r}}(x, s)) \left[\frac{\partial}{\partial s} \int_a^b h(t) \frac{\partial}{\partial t} \Omega_1^{\delta_{0,r}}(s, t | \mu) dt \right] ds \\ & \quad + L'_x(\xi | \psi_1(x, l | \delta_{0,r})) \quad (\psi_1(x, l | \delta_{0,r}) \text{ from (5.12a)}). \end{aligned}$$

In the limit, as $r \rightarrow \infty$ (the $\delta_{0,r}$ being suitably chosen) we get

$$(5.14a) \quad \begin{aligned} & \int_a^b L'_x(\xi | K_1(x, y)) h(y) dy \\ &= \int_{-l}^l d_\mu \int_a^b L'_x(\xi | K_1(x, s)) \left[\frac{\partial}{\partial s} \int_a^b h(t) \frac{\partial}{\partial t} \Omega_1(s, t | \mu) dt \right] ds \\ & \quad + L'_x(\xi | \psi_1(x, l)) \quad (\psi_1(x, l) \text{ from (5.13)}). \end{aligned}$$

In fact, the first member of (5.14a) is obtained as a consequence of (3), (1), (2) and of Theorem 1.2, where we put $g_r(y) = h(y)$ and

$$f_r(y) = L'_x(\xi | K_1^{\delta_{0,r}}(x, y)), \quad \gamma(y) = \gamma(\xi | y).$$

The integral displayed in the 2d member of (5.14a) is obtained from the corre-

sponding integral in (5.14) with the aid of the following considerations. Since

$$\begin{aligned} p_r(s, \mu) &= \frac{\partial}{\partial s} \int_a^b h(t) \frac{\partial}{\partial t} \Omega_1^{\delta_0, r}(s, t | \mu) dt \\ &\rightarrow \frac{\partial}{\partial s} \int_a^b h(t) \frac{\partial}{\partial t} \Omega_1(s, t | \mu) dt = p(s, \mu), \end{aligned}$$

convergence being in the weak sense in s , and since

$$\begin{aligned} L_x(\xi | K_1^{\delta_0, r}(x, s)) &\rightarrow L_x'(\xi | K_1(x, s)) & (\text{as } r \rightarrow \infty), \\ |L_x'(\xi | K_1^{\delta_0, r}(x, s))| &\leq \gamma(\xi | s) \in L_2 & (\text{in } s), \end{aligned}$$

by Theorem 1.5 it is inferred that

$$\begin{aligned} (5.15) \quad q_r(\mu) &= \int_a^b L_x'(\xi | K_1^{\delta_0, r}(x, s)) p_r(s, \mu) ds \\ &\rightarrow \int_a^b L_x'(\xi | K_1(x, s)) p(s, \mu) ds = q(\mu) \quad (\text{as } r \rightarrow \infty). \end{aligned}$$

Moreover, by (3.11a) with

$$g(s) = L_x'(\xi | K_1^{\delta_0, r}(x, s)), \quad \Omega = \Omega_1^{\delta_0, r},$$

and in view of (3), (2), it is concluded that

$$\begin{aligned} (5.15a) \quad \text{Var.}_\mu q_r(\mu) &\leq \left[\int_a^b h^2(x) dx \right]^{1/2} \left[\int_a^b |L_x'(\xi) K_1^{\delta_0, r}(x, s)|^2 ds \right]^{1/2} \\ &\leq \left[\int_a^b h^2(x) dx \right]^{1/2} \left[\int_a^b \gamma^2(\xi | s) ds \right]^{1/2} = A(\xi), \end{aligned}$$

where $A(\xi)$ is independent of r and μ . In consequence of (5.15) and (5.15a), application of Theorem 1.1 is possible, yielding the result

$$\int_{-l}^l dq_r(\mu) \rightarrow \int_{-l}^l dq(\mu),$$

which accounts for the integral in the second member of (5.14a). With $\psi_1(x, l | \delta_{0, r})$ converging weakly (in x , as $r \rightarrow \infty$) to $\psi_1(x, l)$, we have

$$\lim_r L_x'(\xi | \psi_1(x, l | \delta_{0, r})) = L_x'(\xi | \psi_1(x, l))$$

in view of the condition (4). Accordingly, one may consider (5.14a) established.

On letting l in (5.14a) approach infinity, in consequence of (5.13b) it is inferred that (cf. (5.15))

$$\int_a^b L_x'(\xi | K_1(x, y)) h(y) dy = \int_{-\infty}^{\infty} d_{\mu} q(\mu).$$

Accordingly, we observe that (3.22) holds for the class H_m .

In accordance with (3.23) write

$$(5.16) \quad \begin{aligned} \tau_r'(x, l) &= \frac{\partial}{\partial x} \int_{-l}^l d_{\lambda} \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(x, y | \lambda) dy, \\ w_r'(x, l) &= f(x) - \tau_r'(x, l) \end{aligned} \quad (f(x) \in L_2).$$

By (3.23a)–(3.23c)

$$(5.16a) \quad \begin{aligned} \int_a^b \tau_r'^2(x, l) dx &\leq q = \int_a^b f^2(x) dx, \\ w_r'(x, l_{\nu}) &\rightarrow w_r'(x) \quad (\text{as } l_{\nu} \rightarrow \infty), \quad \int_a^b w_r'^2(x) dx \leq 4q, \end{aligned}$$

convergence being in the weak sense; moreover, in view of (3.23d)

$$(5.16b) \quad \int_a^b L_x'(\xi | K_1^{\delta_0, r}(x, y)) w_r'(y) dy = 0.$$

Let $\tau'(x, l)$ be $\tau_r'(x, l)$, with δ_0, r in the integrand of (5.16) deleted, and let $w'(x, l) = f(x) - \tau'(x, l)$. Then because of (3.13), applied with $\alpha(\lambda) = 1$ to Ω_1 , one has

$$(5.16c) \quad \int_a^b \tau'^2(x, l) dx \leq q, *$$

in consequence of which

$$\int_a^b w'^2(x, l) dx \leq 4q.$$

By Theorem 1.3 the latter inequality implies that there exists a subsequence $(0 < l'_1 < l'_2; \lim_{\nu} l'_\nu = \infty)$ such that

$$(5.16d) \quad w'(x, l'_\nu) \rightarrow w'(x) \quad (\text{as } \nu \rightarrow \infty), \quad \int_a^b w'^2(x) dx \leq 4q,$$

convergence being in the *weak sense* (in x). The function $w'(x)$ can also be obtained by a limiting process with the aid of (5.16) and of the last inequality (5.16a); we obtain (cf. Theorem 1.3)

* (5.16c) can also be obtained by a limiting process applied on the basis of (5.16a), with the aid of Theorem 1.3.

(5.16e) $w_r'(x) \rightarrow w'(x)$ [as $r \rightarrow \infty$; weak convergence in x ; suitable $\delta_{0,r}$].

In view of (5.16e) and since (by (2.22') and (2.21'))

$$(5.16f) \quad \begin{aligned} L_x'(\xi | K_1^{\delta_{0,r}}(x, y)) &\rightarrow L_x'(\xi | K_1(x, y)), \\ |L_x'(\xi | K_1^{\delta_{0,r}}(x, y))| &\leq \gamma(\xi | x) \in L_2, \end{aligned}$$

application of Theorem 1.5 to the first member of (5.16b) is possible; thus,

$$(5.17) \quad \int_a^b L_x'(\xi | K_1(x, y)) w'(y) dy = 0.$$

Hence it is observed that the statements previously made in connection with (3.23)–(3.23d) will hold for the class H_m .

If the only solution (included in L_2) of the equation (3.24) [with $L=L'$ and $K=K_1$] is $\phi(y)=0$ (almost everywhere), then in consequence of (5.17), $w'(y)=0$. Now, according to the statement subsequent to (5.16b)

$$w'(x, l'_r) = f(x) - \tau'(x, l'_r);$$

thus, by (5.16d),

$$\tau'(x, l'_r) \rightarrow f(x) \quad (\text{as } r \rightarrow \infty; \text{ in the weak sense}).$$

That is, in view of (5.16) (with $\delta_{0,r}$ deleted)

$$\int_a^x \tau'(x, l'_r) dx = \int_{-l'_r}^{l'_r} d_\lambda \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \rightarrow \int_a^x f(x) dx \quad (\text{as } r \rightarrow \infty).$$

Hence

$$\int_{-\infty}^{\infty} d_\lambda \int_a^b f(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy = \int_a^x f(x) dx. *$$

This formula implies the representation (3.24a), as stated, for the class H_m .

With Δ designating the operation indicated preceding (3.25) it is observed that, by (3.9a) (for $\Omega = \Omega_1$), we have in the sense of weak convergence (in x)

$$\frac{\partial}{\partial x} \Delta \int_{y=a}^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_{0,r}}(x, y | \lambda) dy \rightarrow \frac{\partial}{\partial x} \Delta \int_{y=a}^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy;$$

(3.9) will hold for $\Omega_1^{\delta_{0,r}}$. From (4) it is deduced that

$$(5.18) \quad \begin{aligned} \lim_r L_x \left(\xi \left| \frac{\partial}{\partial x} \Delta \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_{0,r}}(x, y | \lambda) dy \right. \right) \\ = L_x \left(\xi \left| \frac{\partial}{\partial x} \Delta \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \right. \right). \end{aligned}$$

* The integral of the first member is convergent.

It is noted that

$$(5.19) \quad q_r(s) = \frac{\partial}{\partial s} \int_{\lambda'}^{\lambda''} \mu d_\mu \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1^{\delta_0, r}(s, y | \mu) dy \in L_2 \quad (\text{in } s),$$

and that (with $\delta_{0, r}$ suitably chosen)

$$(5.19a) \quad \lim_r q_r(s) = \frac{\partial}{\partial s} \int_{\lambda'}^{\lambda''} \mu d_\mu \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(s, y | \mu) dy = q(s)$$

(weak convergence). In fact, (5.19) follows from (3.13) (with $\alpha(\mu) = \mu$ and $\Omega = \Omega_1^{\delta_0, r}$), while (5.19a) is a consequence of (3.13), (3.12) [with $\alpha(\mu) = \mu$ and $g = 1$ on (a, x) , $g = 0$ on (x, b)] and of Theorem 1.3; we have (cf. (3.13))

$$(5.19b) \quad \int_a^b q^2(s) ds \leq M^2 \int_a^b h^2(s) ds.$$

In virtue of (5.19), (5.19a), and (5.16f) from Theorem 1.5 it is inferred that

$$(5.20) \quad \lim_r \int_a^b L_z'(\xi | K_1^{\delta_0, r}(x, s)) q_r(s) ds = \int_a^b L_z'(\xi | K_1(x, s)) q(s) ds.$$

With (5.18) and (5.20) in view, write (3.25) for $\Omega = \Omega_1^{\delta_0, r}$, $L = L'$, and $K = K_1^{\delta_0, r}$ and pass to the limit, thus obtaining the formula

$$(5.20a) \quad L_z' \left(\xi \left| \frac{\partial}{\partial x} \Delta \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \lambda) dy \right. \right) - \int_a^b L_z'(\xi | K_1(x, s)) q(s) ds = 0;$$

accordingly, it is observed that (3.25) holds for the class H_m , the same being true for (3.25a) (which is obtained from (3.25) by specializing h).

The proof of the important formula (3.26) (for the class H_m) can be effected as follows. In view of (3.22) (with $L = L'$, $K = K_1$, $\Omega = \Omega_1$), (3.26) will hold, for $L = L'$, $K = K_1$, $\Omega = \Omega_1$, provided

$$(5.21) \quad \int_{-\infty}^{\infty} d_\mu \int_a^b L_z'(\xi | K_1(x, s)) \left[\frac{\partial}{\partial s} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(s, y | \mu) dy \right] ds = \int_{-\infty}^{\infty} \frac{1}{\mu} d_\mu L_z' \left(\xi \left| \frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \mu) dy \right. \right) = N(\xi).$$

On writing

$$(5.21a) \quad N_l(\xi) = \int_{-l}^l \frac{1}{\mu} d_\mu L_z' \left(\xi \left| \frac{\partial}{\partial x} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \mu) dy \right. \right),$$

and

$$-l = \lambda_0 < \lambda_1 < \dots < \lambda_{m_1} = l \quad (\Delta_r \text{ corresponding to } (\lambda_{r-1}, \lambda_r)),$$

we have

$$(5.21b) \quad N_l(\xi) = \lim_{m_1} N_{l, m_1}(\xi),$$

where

$$(5.21c) \quad N_{l, m_1}(\xi) = \sum_{r=1}^{m_1} \frac{1}{\mu_r} L'_z \left(\xi \left| \frac{\partial}{\partial x} \Delta_r \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(x, y | \mu) dy \right. \right)$$

$(\mu_r \text{ in } (\lambda_{r-1}, \lambda_r))$. By (3.25) (for L' , K_1 , Ω_1) and (5.21c)

$$(5.21d) \quad N_{l, m_1}(\xi) = \sum_{r=1}^{m_1} \frac{1}{\mu_r} \int_a^b L'_z(\xi | K_1(x, s)) \cdot \left[\frac{\partial}{\partial s} \int_{\lambda_{r-1}}^{\lambda_r} \mu d\mu \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(s, y | \mu) dy \right] ds.$$

Applying to the integral displayed in (5.21d) the first identity (3.14), with

$$g(s) = L'_z(\xi | K_1(x, s)), \quad \alpha(\mu) = \mu, \quad \Omega = \Omega_1,$$

it is concluded that (5.21d) may be written in the form

$$(5.21e) \quad N_{l, m_1}(\xi) = \sum_{r=1}^{m_1} \frac{1}{\mu_r} \int_{\lambda_{r-1}}^{\lambda_r} \mu \cdot d\mu \left\{ \int_a^b L'_z(\xi | K_1(x, s)) \left[\frac{\partial}{\partial s} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(s, y | \mu) dy \right] ds \right\}.$$

Thus

$$(5.21f) \quad N_{l, m_1}(\xi) = \sum_{r=1}^{m_1} \frac{1}{\mu_r} \Delta_r \gamma(\mu), \quad \gamma(\mu) = \int_{-l}^{\mu} \mu d\mu \{ \dots \},$$

and, in view of (5.21b),

$$(5.22) \quad N_l(\xi) = \int_{-l}^l \frac{1}{\mu} d\mu \gamma(\mu).$$

By definition of $\gamma(\mu)$ (cf. (5.21f)), with $\{ \dots \}$ from (5.21e), we have

$$(5.22a) \quad \begin{aligned} N_l(\xi) &= \int_{-l}^l \frac{1}{\mu} [\mu d\mu \{ \dots \}] \\ &= \int_{-l}^l d\mu \left\{ \int_a^b L'_z(\xi | K_1(x, s)) \left[\frac{\partial}{\partial s} \int_a^b h(y) \frac{\partial}{\partial y} \Omega_1(s, y | \mu) dy \right] ds \right\}. \end{aligned}$$

By (5.22a), (5.21a)

$$\lim_l N_l(\xi) = N(\xi) = \text{first member of (5.21)}.$$

Whence it is observed that (5.21) and consequently (3.26) (for L' , K_1 , Ω_1) have been established.

We shall now proceed to prove the statement in connection with (3.27) for the class H_m . The identity to be proved is

$$\begin{aligned} L'_z \left(\xi \left| \frac{\partial}{\partial x} \int_{\lambda'}^{\lambda''} \mu d_\mu \Omega_1(x, y | \mu) \right. \right) \\ (5.23) \quad - \lambda \int_a^b L'_z(\xi | K_1(x, s)) \left[\frac{\partial}{\partial s} \int_{\lambda'}^{\lambda''} \mu d_\mu \Omega_1(s, y | \mu) \right] ds \\ = L'_z \left(\xi \left| \frac{\partial}{\partial x} \int_{\lambda'}^{\lambda''} (\mu - \lambda) d_\mu \Omega_1(x, y | \mu) \right. \right). \end{aligned}$$

In view of (3.25a) (for L' , K_1 , Ω_1), (5.23a) will hold if

$$\begin{aligned} L'_z \left(\xi \left| \frac{\partial}{\partial x} \int_{\lambda'}^{\lambda''} \mu d_\mu \Omega_1(x, y | \mu) \right. \right) - \lambda L'_z \left(\xi \left| \frac{\partial}{\partial x} \Delta \Omega_1(x, y | \lambda) \right. \right) \\ = L'_z \left(\xi \left| \frac{\partial}{\partial x} \int_{\lambda'}^{\lambda''} (\mu - \lambda) d_\mu \Omega_1(x, y | \mu) \right. \right); \end{aligned}$$

that is, if

$$(5.23a) \quad L'_z \left(\xi \left| \frac{\partial}{\partial x} \Delta \Omega_1(x, y | \lambda) \right. \right) = L'_z \left(\xi \left| \frac{\partial}{\partial x} \int_{\lambda'}^{\lambda''} d_\mu \Omega_1(x, y | \mu) \right. \right).$$

Now, (5.23a) holds since

$$\Delta \Omega_1(x, y | \lambda) = \int_{\lambda'}^{\lambda''} d_\mu \Omega_1(x, y | \mu).$$

Thus, the result previously stated with respect to (3.27) holds for the class H_m .

The developments of this section may be summed in the theorem:

THEOREM 5.1. *Suppose classes H_n are specified by Definition 2.2. For every finite m (>0) the following will hold. If $K(x, y) \in H_m$ and if "associated" with $K(x, y)$ there is an operator L (cf. Definition 2.3), then the statements made in §3 will hold true with respect to this kernel and this operator.*

6. Kernels of class H_1 . For kernels of class H_1 with which operators L (Definition 2.3) can be "associated," as remarked by Carleman many results

of (C, chap. 2) can be extended. In view of the purpose to examine the possibility of such extension to our classes H_n it will be essential to investigate in some detail the situation with respect to H_1 .

Thus, suppose $K(x, y) \in H_1$, an operator L being "associated" with $K(x, y)$.

We write down the equations

$$(6.1) \quad \phi^{\delta_0}(x) - \lambda \int_a^b K^{\delta_0}(x, y) \phi^{\delta_0}(y) dy = f(x),$$

$$(6.1a) \quad L_x(\xi | \phi^{\delta_0}(x)) - \lambda \int_a^b L_x(\xi | K^{\delta_0}(x, y)) \phi^{\delta_0}(y) dy = L_x(\xi | f(x)),$$

$$(6.2) \quad L_x(\xi | \phi(x)) - \lambda \int_a^b L_x(\xi | K(x, y)) \phi(y) dy = L_x(\xi | f(x)),$$

$$(6.2a) \quad L_x(\xi | \phi(x)) - \lambda \int_a^b L_x(\xi | K(x, y)) \phi(y) dy = 0.$$

As indicated in (C), if $\phi^{\delta_0, r}(x)$ is a solution of (6.1) (for $\delta = \delta_0, r$ suitably chosen; $r = 1, 2, \dots$; $\lim \delta_0, r = 0$), then $\phi^{\delta_0, r}(x) \rightarrow \phi(x)$ (weakly in x) and (3.19a) will hold; moreover, $\phi(x)$ will be a solution of (6.2). On writing, conforming with (C),

$$(6.3) \quad \phi^{\delta_0}(x) = \frac{i\lambda'}{2\beta} [f(x) + \psi^{\delta_0}(x)], \quad \phi(x) = \frac{i\lambda'}{2\beta} [f(x) + \psi(x)],$$

where λ' is the conjugate of λ , we conclude that

$$(6.3a) \quad \int_a^b |\psi^{\delta_0, r}(x)|^2 dx = \int_a^b |f(x)|^2 dx$$

and, in the limit,

$$(6.3b) \quad \int_a^b |\psi(x)|^2 dx \leq \int_a^b |f(x)|^2 dx.$$

It is observed that (6.3a) can be established with the aid of the relation, found in (C),

$$(6.3c) \quad \begin{aligned} \frac{-\beta}{|\lambda|^2} \int_a^b |\phi^{\delta_0}(x)|^2 dx &= \frac{1}{2i\lambda} \int_a^b f(x) \phi_1^{\delta_0}(x) dx \\ &\quad - \frac{1}{2i\lambda'} \int_a^b \bar{f}(x) \phi^{\delta_0}(x) dx. \end{aligned}$$

Here and in the sequel, ϕ_1 denotes the conjugate of ϕ .

We shall prove the following fact. In order that (6.3b) should hold (when

$\lambda = \beta \neq 0$) with the equality sign it is necessary and sufficient that

$$(6.4) \quad \lim \int_a^b |\phi^{b_0, r}(x)|^2 dx = \int_a^b |\phi(x)|^2 dx.$$

In fact, it is noted that

$$\lim \int_a^b f(x) \phi_1^{b_0, r}(x) dx = \int_a^b f(x) \phi(x) dx,$$

inasmuch as $f(x) \in L_2$ and other conditions of Theorem 1.4 hold. Thus, by (6.3c)

$$(6.4a) \quad \begin{aligned} \frac{-\beta}{|\lambda|^2} \lim \int_a^b |\phi^{b_0, r}(x)|^2 dx &= \frac{1}{2i\lambda} \int_a^b f(x) \phi_1(x) dx \\ &\quad - \frac{1}{2i\lambda'} \int_a^b \bar{f}(x) \phi(x) dx = \gamma, \end{aligned}$$

and, if (6.4) holds, there will be on hand an equality like (6.3c) with $\phi^{b_0}(x)$ replaced by $\phi(x)$; from this relation with the aid of the second one of (6.3) we obtain (6.3b) with the equality sign. Thus (6.4) is a sufficient condition. If, on the other hand, (6.4) does not hold, it is observed that, inasmuch as the limit in (6.4a) exists,

$$(6.4b) \quad \lim \int_a^b |\phi^{b_0, r}(x)|^2 dx = \frac{\gamma |\lambda|^2}{-\beta} = \left| \frac{\gamma}{\beta} \right| |\lambda|^2 > \int_a^b |\phi(x)|^2 dx;$$

this inequality follows by a theorem of F. Riesz according to which

$$\limsup \int_a^b |f_r(x)|^2 dx \geq \int_a^b |f(x)|^2 dx,$$

whenever $f_r(x) \in L_2$ and $f_r(x) \rightarrow f(x)$ (in the weak sense). Now, in consequence of (6.4b)

$$(6.4c) \quad \left| \frac{\beta}{\lambda|^2} \int_a^b |\phi(x)|^2 dx < |\gamma| \quad (\gamma \text{ from (6.4a)});$$

substitution of $\phi(x)$ from (6.3) and (6.4c) will result in (6.3b), with the inequality sign. Hence the statement in connection with (6.4) is seen to be true.*

Of special interest appear to be operators L , which in addition to the condi-

* The corresponding result in (C) depends on the possibility of interchange of order of integration in a certain double integral.

tions of Definition 2.3 satisfy the following. L is defined for ξ in a set Γ dense in itself; moreover, for ξ_1, ξ_2 , in Γ ,

$$(6.5) \quad \int_a^b |L_x(\xi_1 | K^{\delta_0}(x, y)) - L_x(\xi_2 | K^{\delta_0}(x, y))|^2 dy \leq G(\xi_1, \xi_2),$$

where $G(\xi_1, \xi_2)$ is independent of δ_0 and

$$(6.5a) \quad G(\xi_1, \xi_2) \rightarrow 0 \quad (\text{as } \xi_1 - \xi_2 \rightarrow 0; \xi_1, \xi_2 \text{ in } \Gamma).$$

In consequence of (2.22) the limit of the integrand (for $\delta_0 \rightarrow 0$) in (6.5) exists; by (2.21) the integrand is less than

$$[\gamma(\xi_1 | y) + \gamma(\xi_2 | y)]^2 \leq L_1 \quad (\text{in } y)$$

for ξ_1, ξ_2 on Γ . Hence by passage to the limit, from (6.5) it is inferred that

$$(6.5b) \quad \int_a^b |L_x(\xi_1 | K(x, y)) - L_x(\xi_2 | K(x, y))|^2 dy \leq G(\xi_1, \xi_2).$$

Let $\phi(x)$ be a solution of (6.2) obtained by a limiting process as indicated subsequent to (6.2a). Then, by (6.2) and by the inequality of Schwartz,

$$\begin{aligned} & |L_x(\xi_1 | \phi(x) - f(x)) - L_x(\xi_2 | \phi(x) - f(x))|^2 \\ &= |\lambda|^2 \left| \int_a^b [L_x(\xi_1 | K(x, y)) - L_x(\xi_2 | K(x, y))] \phi(y) dy \right|^2 \\ &\leq |\lambda|^2 \int_a^b |\phi(y)|^2 dy \int_a^b |L_x(\xi_1 | K(x, y)) - L_x(\xi_2 | K(x, y))|^2 dy. \end{aligned}$$

If (6.5) and (6.5a) hold, then in consequence of (6.5b) and of (3.19a)

$$(6.6) \quad \begin{aligned} & |L_x(\xi_1 | \phi(x) - f(x)) - L_x(\xi_2 | \phi(x) - f(x))|^2 \\ &\leq \frac{|\lambda|^4}{\beta^2} \int_a^b |f(x)|^2 dx G(\xi_1, \xi_2) \end{aligned}$$

(for ξ_1, ξ_2 on Γ); thus, under (6.5) and (6.5a), for every solution $\phi(x)$, included in L_2 , of (6.2) the function

$$(6.6a) \quad L_x(\xi | \phi(x) - f(x))$$

will be continuous in ξ for ξ on Γ ($\beta \neq 0$).

When $\phi^{\delta_0}(x)$ satisfies (6.1), in consequence of (6.1a) one has

$$\begin{aligned} |L_x(\xi | \phi^{\delta_0}(x))| &\leq |\lambda| \left[\int_a^b |L_x(\xi | K^{\delta_0, r}(x, y))|^2 dy \right]^{1/2} \left[\int_a^b |\phi^{\delta_0}(x)|^2 dx \right]^{1/2} \\ &\quad + |L_x(\xi | f(x))|, \end{aligned}$$

whence, by (3.19a) for $(\phi^{b_0}(x))$ and (2.21) (for $n=1$),

$$(6.6b) \quad |L_x(\xi | \phi^{b_0}(x))| \leq \frac{|\lambda|^2}{|\beta|} \left[\int_a^b |f(x)|^2 dx \right]^{1/2} \left[\int_a^b \gamma^2(\xi | y) dy \right]^{1/2} + |L_x(\xi | f(x))|.$$

Similarly

$$(6.6c) \quad |L_x(\xi | \phi^{b_0}(x) - f(x))|^2 \leq \frac{|\lambda|^4}{\beta^2} \int_a^b |f(x)|^2 dx \int_a^b \gamma^2(\xi | y) dy.$$

On the other hand, by (6.1a) (for ξ_1 and ξ_2) we obtain the inequality subsequent to (6.5b), with $\phi(x)$ and $K(x, y)$ replaced by $\phi^{b_0}(x)$ and $K^{b_0}(x, y)$, respectively; an application of (3.19a) (for $\phi^{b_0}(x)$) will yield

$$(6.7) \quad |L_x(\xi_1 | \phi^{b_0}(x) - f(x)) - L_x(\xi_2 | \phi^{b_0}(x) - f(x))|^2 \leq \frac{|\lambda|^4}{\beta^2} \int_a^b |f(x)|^2 dx G(\xi_1, \xi_2),$$

(for ξ_1, ξ_2 on Γ) if L is such that (6.5) holds.

If (6.5) and (6.5a) are assumed, in view of (6.6b) and of (6.7) application of Vitali's theorem (on limits of analytic functions) is possible in a manner analogous to that in (C, p. 55). The following result is obtained.

Let $K(x, y) \in H_1$ and L be an "associated" operator (Definition 2.3), satisfying (6.5), (6.5a) and let $\Gamma\lambda = \beta \neq 0$. For a suitable choice of $\delta_{0,r}$ not only will $\phi^{b_{0,r}}(x)$ converge (weakly) to a solution $\phi(x)$ satisfying (6.2) but the function $L_x(\xi | \phi(x) - f(x))$ will be continuous in ξ (for ξ in Γ) and will be regular in λ for all non-real λ (when ξ is in Γ).

The analogue of (C, Theorems III₂, IV₂), for kernels $K(x, y) \in H_1$ and having "associated" with them an operator L , is obtained by passage to the limit. The result reads as follows. Given a value $\lambda = \lambda_0$ ($\lambda_0 = \alpha_0 + i\beta_0$; $\beta_0 \neq 0$), there exists an operator T_0 (depending on λ_0 , but independent of f) so that

$$(6.8) \quad \phi(x) = T_0(f(x))$$

will constitute a solution of (6.2); moreover,

$$(6.8a) \quad \int_a^b T_0(f_2(x)) \cdot f_1(x) dx = \int_a^b T_0(f_1(x)) \cdot f_2(x) dx,$$

whenever $f_1(x), f_2(x) \in L_2$. In particular, if for $\lambda = \lambda_0$ the equation (6.2) has only one solution $\phi(x) \in L_2$, (6.8), (6.8a) will hold.

The analogue to (C, Theorem II₂) will be as follows.

If the operator L is such that $L_x(\xi | q(x))$ is real for $q(x)$ real and the

$$(6.9) \quad \text{conjugate of } L_x(\xi | q(x)) = L_x(\xi | \bar{q}(x)),$$

and if for a particular λ_0 ($\Gamma\lambda_0 = \beta_0 \neq 0$) the homogeneous equation (6.2a) has no solutions included in L_2 , except $\phi(x) = 0$ (almost everywhere), then for all non-real values of λ (6.2a) will have no solutions included in L_2 , except zero (almost everywhere).

The proof of this theorem is closely analogous to that of (C, Theorem II₂). However, in view of the extension, to be given in the sequel, of this result to classes H_n it is desirable to outline briefly a sketch of the proof.

If the theorem is not true, then

$$(6.10) \quad L_x(\xi | \phi(x)) - \lambda_0 \int_a^b L_x(\xi | K(x, y))\phi(y)dy = L_x(\xi | (1 - \lambda_0/\lambda)\phi(x)),$$

where $\phi(x) \in L_2$, $\phi(x) \neq 0$, and $\phi(x)$ is a solution of (6.2a) for a value λ ($\Gamma\lambda = \beta \neq 0$). Using (6.9) one obtains

$$(6.10a) \quad L_x(\xi | \phi_1(x)) - \lambda_0 \int_a^b L_x(\xi | K(x, y))\phi_1(y)dy = L_x(\xi | (1 - \lambda_0/\lambda')\phi_1(x)).$$

In consequence of the statement in connection with (6.8) and (6.8a),

$$(6.11) \quad \left(1 - \frac{\lambda_0}{\lambda'}\right) \int_a^b |\phi(x)|^2 dx = \left(1 - \frac{\lambda_0}{\lambda}\right) \int_a^b |\phi(x)|^2 dx,$$

and, inasmuch as $\int |\phi|^2 dx \neq 0$, necessarily $\beta = 0$ which is contrary to hypothesis.

Similarly, following the lines indicated in (C) one may prove the following analogue to the important result of (C, Theorem V₂).

If the operator L (associated with $K(x, y) \in H_1$) satisfies the condition (6.9), then the number of linearly independent solutions [included in L_2] of the homogeneous equation (6.2a) is the same for all λ ($\Gamma\lambda = \beta \neq 0$).

With respect to linear independence of solutions of (6.2a) the following will hold.

Let $\phi_1(x), \dots, \phi_n(x)$ be solutions included in L_2 of the homogeneous equation (6.2a), corresponding to the distinct values of λ ,

$$\lambda_1, \dots, \lambda_n \quad (\lambda_v \neq 0; v = 1, \dots, n);$$

the $\phi_j(x)$ will be linearly independent if L ("associated" with $K(x, y)$) is such that

$$(6.12) \quad L_x(\xi | q(x)) = 0 \quad (q(x) \in L_2)$$

implies that $q(x) = 0$ (almost everywhere).

In fact, if this theorem is not true, then for some c_r

$$c_1\phi_1(y) + \dots + c_n\phi_n(y) = 0 \quad (\text{not all } c_r = 0).$$

Multiplying by $L_x(\xi|K(x, y))dy$, integrating and making use of

$$(6.13) \quad L_x(\xi|\phi_r(x)) = \lambda_r \int_a^b L_x(\xi|K(x, y))\phi_r(y)dy,$$

we obtain

$$(6.13a) \quad \sum_{r=1}^n \frac{c_r}{\lambda_r} L_x(\xi|\phi_r(x)) = L_x\left(\xi \left| \sum_{r=1}^n \frac{c_r}{\lambda_r} \phi_r(x) \right. \right) = 0.$$

In view of the property in connection with (6.12),

$$\sum_{r=1}^n \frac{c_r}{\lambda_r} \phi_r(x) = 0.$$

Repeating this process a number of times and at each step making use of (6.13) and of the property referred to above, a set of equations is obtained which cannot be satisfied, unless all the c_r are zero.

Let us examine now the question of the range of values which, for x and $\lambda = \lambda_1$ ($\beta_1 \neq 0$) fixed, could be assumed by the solutions of (6.2).*

According to the italicized statement subsequent to (6.11) the number of linearly independent solutions of (6.2a) (where $K(x, y) \in H_1$ and L is an associated operator) is the same for all non-real λ . It can be arranged to have these solutions forming an orthogonal and normal set (for a fixed λ). Let $\Phi_1(x), \Phi_2(x), \dots$ constitute a full set of such description for λ_1 . Let $\phi(x)$ be any solution of (6.2) for λ_1 , the corresponding $\psi(x)$ being defined (cf. (6.3)) by

$$(6.14) \quad \phi(x) = \frac{i\lambda_1'}{2\beta_1} (f(x) + \psi(x)) \quad (\lambda_1' = \text{conjugate of } \lambda_1).$$

Then the result of the same form as given in (C, pp. 71, 72) will hold for equations (6.2):

$$(6.15) \quad |\phi(x) - c_r(x, \lambda_1)| \leq r(x, \lambda_1) [c(x, \lambda_1) = (i\lambda_1'/2\beta_1)(f(x) + w(x))],$$

$$(6.15a) \quad r^2(x, \lambda_1) = \frac{|\lambda_1|^2}{4\beta_1^2} \int_a^b [|f(t)|^2 - |w(t)|^2] dt \sum_r |\Phi_r(x)|^2,$$

where $w(x)$ is from the Fourier-expansion (in terms of the $\Phi_r(x)$) of $\psi(x)$,

$$(6.15b) \quad \psi(x) = w(x) + \sum c_r \Phi_r(x)$$

* A problem of this type is treated in (C) for kernels not of class H_1 and without the aid of operators L .

(cf. (6.14)), and is independent of $\phi(x)$.^{*} To establish this one needs only to take note of the inequality (6.3b) and to follow the procedure indicated in (C).

Corresponding to the function $w(x)$ (of (6.15b)) there is a particular solution (for $\lambda_1; \beta \neq 0$) of (6.2),

$$(6.16) \quad \phi_0(x) = \frac{i\lambda_1}{2\beta_1} (f(x) + w(x)).$$

In view of the statement in connection with (6.4), from (6.15) and (6.15a) it is inferred that, if there exists a sequence $\{\phi^{\delta_0, r}(x)\}$ † which converges in the weak sense to $\phi_0(x)$, while

$$(6.16a) \quad \lim \int_a^b |\phi^{\delta_0, r}(x)|^2 dx = \int_a^b |\phi_0(x)|^2 dx$$

(cf. (6.16)), then $\phi_0(x)$ is the only solution (for λ_1) included in L_2 of (6.2).

Consider a value $\lambda = \lambda_1$ (with $\beta_1 \neq 0$). If not every solution of the homogeneous equation (6.2a), for $\lambda = \lambda_1$ is zero (almost everywhere), then the number $r(x, \lambda_1)$, involved in (6.15a), will be distinct from zero at least for some $f(x) \in L_2$, provided the operator L satisfies the condition (6.9). In fact, with the aid of the latter condition the procedure given in (C) (for the demonstration of an analogous result) is applicable, leading to the stated assertion.

7. Extension of the results of §6 to the classes H_n . With a view to proof by induction let us assume that the following holds for kernels $K(x, y)$ of classes H_n ($n = 1, 2, \dots, m-1$), it being understood that "associated" with every kernel $K(x, y)$, under consideration, there is an operator L (Definition 2.3). For convenience we collect the requisite equations:

$$(7.1) \quad \phi^{\delta_0, \dots, \delta_{n-1}}(x) - \lambda \int_a^b K^{\delta_0, \dots, \delta_{n-1}}(x, y) \phi^{\delta_0, \dots, \delta_{n-1}}(y) dy = f(x),$$

$$(7.1a) \quad L_x(\xi | \phi^{\delta_0, \dots, \delta_{n-1}}(x)) - \lambda \int_a^b L_x(\xi | K^{\delta_0, \dots, \delta_{n-1}}(x, y)) \phi^{\delta_0, \dots, \delta_{n-1}}(y) dy \\ = L_x(\xi | f(x)),$$

$$(7.2) \quad L_x(\xi | \phi(x)) - \lambda \int_a^b L_x(\xi | K(x, y)) \phi(y) dy = L_x(\xi | f(x)),$$

$$(7.2a) \quad L_x(\xi | \phi(x)) - \lambda \int_a^b L_x(\xi | K(x, y)) \phi(y) dy = 0.$$

If a solution $\phi(x)$ (for a value λ , with $\beta \neq 0$) of (7.2) is the repeated limit in the weak sense of solutions of (7.1) (cf. (3.19)), then

^{*} $w(x)$ is the analogue of $\psi_0(x)$ of (C, p. 71); that is, $w(x)$ minimizes $\int (\psi')^2 dx$ (for $\phi(x)$ satisfying (6.2)). Thus, $w(x)$ is a particular function $\psi(x)$. Obviously $\int w \Phi_y dx = 0$.

† $\phi^{\delta_0, r}(x)$ a solution of (6.1) (for $\delta_0 = \delta_{0, r}$).

$$(7.3) \quad \int_a^b |\psi(x)|^2 dx \leq \int_a^b |f(x)|^2 dx \quad [\phi(x) = (i\lambda'/2\beta)(f(x) + \psi(x))].$$

The necessary and sufficient condition under which (7.3) holds with the equality sign is that

$$(7.4) \quad \lim_{\delta_0, r} \lim_{\delta_1, r} \cdots \lim_{\delta_{n-1}, r} \int_a^b |\phi^{\delta_0, r, \delta_1, r, \dots, \delta_{n-1}, r}(x)|^2 dx = \int_a^b |\phi(x)|^2 dx$$

(the $\delta_{v, r}$ from (3.19)).

If the operator L ("associated" with $K(x, y)$) is defined for ξ in a set Γ , dense in itself, and if

$$(7.5) \quad \int_a^b |L_x(\xi_1 | K^{\delta_0, \dots, \delta_{n-1}}(x, y)) - L_x(\xi_2 | K^{\delta_0, \dots, \delta_{n-1}}(x, y))|^2 dy \leq G(\xi_1, \xi_2) \\ [G(\xi_1, \xi_2) \rightarrow 0 \text{ as } \xi_1 - \xi_2 \rightarrow 0; \xi_1, \xi_2 \text{ in } \Gamma],$$

then (7.5) will hold also for $K(x, y)$.

Every solution $\phi(x) \in L_2$ of (7.2) is such that

$$(7.5a) \quad L_x(\xi | \phi(x) - f(x))$$

is continuous in ξ for ξ in Γ , provided (7.5) holds.

Let a solution $\phi(x)$ of (7.2) be defined by the repeated limiting process of (3.19). If (7.5) holds, then with a suitable choice of the $\delta_{v, r}$,

$$(7.5b) \quad L_x(\xi | \phi(x) - f(x))$$

will be continuous in ξ , for ξ in Γ , and will be regular in λ for all non-real λ (ξ in Γ).

(1) *The statement in connection with (6.8), (6.8a) holds with respect to the nonhomogeneous equation (7.2) for kernels of classes H_v ($v < m$).*

Also the following holds.

(2) *If $L_x(\xi | q(x))$ is real for $q(x)$ real and (6.9) holds and if the equation (7.2a) has no solutions included in L_2 , except zero, then the same will be true for all non-real values of λ ; the number of linearly independent solutions, included in L_2 , of (7.2a) is the same for all non-real λ .*

(3) *Regarding linear independence we have the result, previously stated in connection with (6.12), holding for classes H_1, \dots, H_{m-1} .*

(4) *The result stated in connection with (6.14)–(6.15b) holds for H_n ($n < m$).*

If $\phi_0(x)$ is a solution for λ_1 ($\beta_1 \neq 0$) of (7.2), which is a repeated limit (in the weak sense) as indicated in (3.19) and which is such that

$$(7.6) \quad w(x) = (2\beta_1/i\lambda_1') (\phi_0(x) - f(x))$$

renders

$$(7.6a) \quad \int_a^b |\psi(x)|^2 dx$$

$[\phi(x) = (i\lambda_1'/2\beta_1)(f(x) + \psi(x))$ solutions for λ_1 of (7.2)] minimum, then $\phi_0(x)$ is the only solution included in L_2 of (7.2), provided (7.4) holds (with $\phi(x) = \phi_0(x)$).

(5) *The italicized statement at the end of §6 holds with respect to the homogeneous equation (7.2a) for classes H_1, \dots, H_{m-1} .*

All of the above properties have been verified in §6 for kernels of class H_1 (with "associated" operators L). We shall now establish these properties for H_m . Let $K^1(x, y) \in H_m$ and let L^1 be an "associated" operator. Then $K^{1\delta_0}(x, y) \in H_{m-1}$; moreover, as indicated at the beginning of §5, L^1 will be also associated with $K^{1\delta_0}(x_1)$.

By (7.3), applied to $K^{1\delta_0}(x, y)$,

$$(7.7) \quad \int_a^b |\psi^{\delta_0}(x)|^2 dx \leq \int_a^b |f(x)|^2 dx.$$

With $\phi^1(x) = \lim \phi^{\delta_0, r}(x)$ (in the weak sense) we shall have $\psi^1(x) = \lim \psi^{\delta_0, r}(x)$ (in the weak sense), where

$$\phi^{\delta_0}(x) = i\lambda'/2\beta(f(x) + \psi^{\delta_0}(x)), \quad \phi'(x) = i\lambda'/2\beta(f(x) + \psi^1(x)).$$

In view of the theorem of Riesz, stated subsequent to (6.4b),

$$(7.7a) \quad \limsup \int_a^b |\psi^{\delta_0, r}(x)|^2 dx \geq \int_a^b |\psi^1(x)|^2 dx.$$

By (7.7) and (7.7a)

$$(7.7b) \quad \int_a^b |\psi^1(x)|^2 dx \leq \int_a^b |f(x)|^2 dx,$$

which is (7.3) for the class H_m .

Since $K^{\delta_0, \dots, \delta_{m-1}}(x, y) \in L_2$ (in x, y) the identity (6.3c) (for $\phi^{\delta_0, \dots, \delta_{m-1}}(x)$) will hold:

$$(7.8) \quad \frac{-\beta}{|\lambda|^2} \int_a^b |\phi^{\delta_0, \dots, \delta_{m-1}}(x)|^2 dx = \frac{1}{2i\lambda} \int_a^b f(x) \phi_1^{\delta_0, \dots, \delta_{m-1}}(x) dx \\ - \frac{1}{2i\lambda'} \int_a^b \bar{f}(x) \phi^{\delta_0, \dots, \delta_{m-1}}(x) dx.$$

Suppose (7.4) holds for ϕ^1 ; thus

$$(7.9) \quad \lim_{\delta_{0,r}} \lim_{\delta_{1,r}} \cdots \lim_{\delta_{m-1,r}} \int_a^b |\phi^{\delta_{0,r}, \dots, \delta_{m-1,r}}(x)|^2 dx = \int_a^b |\phi^1(x)|^2 dx.$$

Now (cf. (3.19) for ϕ^1)

$$\begin{aligned} \lim_{\delta_{m-1,r}} \phi^{\delta_{0,r}, \dots, \delta_{m-1,r}}(x) &= \phi^{\delta_{0,r}, \dots, \delta_{m-2,r}}(x), \\ \lim_{\delta_{m-2,r}} \phi^{\delta_{0,r}, \dots, \delta_{m-2,r}}(x) &= \phi^{\delta_{0,r}, \dots, \delta_{m-3,r}}(x), \dots, \lim_{\delta_{0,r}} \phi^{\delta_{0,r}}(x) = \phi^1(x) \end{aligned}$$

[in the sense of weak convergence], with the functions involved included in L_2 ; hence repeated application of Theorem 1.4 to the second member of (7.8) is possible yielding the result (when $\delta_r = \delta_{r,r}$)

$$\begin{aligned} \lim_{\delta_{0,r}} \cdots \lim_{\delta_{m-1,r}} \left[\frac{1}{2i\lambda} \int_a^b f(x) \phi^{\delta_{0,r}, \dots, \delta_{m-1,r}}(x) dx - \frac{1}{2i\lambda} \int_a^b \bar{f}(x) \phi^{\delta_{0,r}, \dots, \delta_{m-1,r}}(x) dx \right] \\ = \frac{1}{2i\lambda} \int_a^b f(x) \phi^1(x) dx - \frac{1}{2i\lambda'} \int_a^b \bar{f}(x) \phi^1(x) dx. \end{aligned}$$

This, together with (7.9) implies

$$(7.9a) \quad \frac{-\beta}{|\lambda|^2} \int_a^b |\phi^1(x)|^2 dx = \frac{1}{2i\lambda} \int_a^b f(x) \phi^1(x) dx - \frac{1}{2i\lambda'} \int_a^b \bar{f}(x) \phi^1(x) dx.$$

Substituting in (7.9a) $\phi^1(x)$ in terms of $\psi^1(x)$ we obtain the equality

$$(7.9b) \quad \int_a^b |\psi^1(x)|^2 dx = \int_a^b |f(x)|^2 dx,$$

which is observed to be a consequence of (7.9). Suppose now that (7.9) does not hold. Since the repeated limit displayed preceding (7.9a) exists even if (7.9) does not hold, in consequence of (7.8) it can be asserted that

$$(7.10) \quad \begin{aligned} \lim_{\delta_{0,r}} \cdots \lim_{\delta_{m-1,r}} \int_a^b |\phi^{\delta_{0,r}, \dots, \delta_{m-1,r}}(x)|^2 dx &= \gamma \\ &= \left[\frac{1}{2i\lambda} \int_a^b f(x) \phi^1(x) dx - \frac{1}{2i\lambda'} \int_a^b \bar{f}(x) \phi^1(x) dx \right] \frac{|\lambda|^2}{-\beta}. \end{aligned}$$

By the theorem of Riesz (text subsequent to (6.4b)), (7.10) will imply

$$\gamma \geq \int_a^b |\phi^1(x)|^2 dx.$$

This, in view of our previous assumption that (7.9) does not hold, yields the inequality

$$(7.10a) \quad \gamma > \int_a^b |\phi^1(x)|^2 dx.$$

Substituting in (7.10a) the expression for γ from the last member of (7.10) and replacing $\phi^1(x)$ in terms of $\psi^1(x)$, we infer that failure of (7.9) to hold implies

$$(7.10b) \quad \int_a^b |\psi^1(x)|^2 dx < \int_a^b |f(x)|^2 dx.$$

Accordingly, the statement with respect to (7.4) holds for the class H_m .

The property stated with reference to (7.5) will hold for $K^1(x, y)$ and L^1 in consequence of (2.22), of (2.21) (applied to the kernel in question) and of Lebesgue's theorem on passage to the limit under the integral sign (we keep $\beta \neq 0$).

Under (7.5) (for $K^{1\delta_0, \dots, \delta_{m-1}}$ and L^1) the function $L_2^1(\xi | \phi^1(x) - f(x))$, where $\phi^1(x)$ is any solution included in L_2 of (7.2) [that is, of (7.2) with L^1 and K^1], will be continuous in ξ (ξ in Γ). In fact, by the same method as used before we obtain the inequality (6.6) for L^1 and ϕ^1 , which justifies the above assertion.

If $\phi^1(x)$ is a repeated limit in accordance with (3.19), then

$$(7.11) \quad \phi^{\delta_0, r}(x) \rightarrow \phi^1(x) \quad (\text{weakly}),$$

where

$$\phi^{\delta_0, r}(x) = \lim_{\delta_{1, r}} \dots \lim_{\delta_{m-1, r}} \phi^{\delta_0, r, \delta_{1, r}, \dots, \delta_{m-1, r}}(x) \quad (\text{weakly}),$$

and $\phi^{\delta_0, r}(x)$ satisfies

$$(7.11a) \quad L_2^1(\xi | \phi^{\delta_0, r}(x)) - \lambda \int_a^b L_2^1(\xi | K^{\delta_0, r}(x, y)) \phi^{\delta_0, r}(y) dy = L_2^1(\xi | f(x)),$$

with $K^{\delta_0, r}$ formed corresponding to $K^1(x, y)$. Proceeding with respect to (7.11a) as before one obtains the inequalities (6.6b), (6.6c), (6.7) (for L^1); the latter inequality will hold under (7.5) (for $L^1, K^{1\delta_0, \dots, \delta_{m-1}}$). With the aid of Vitali's theorem these inequalities enable us to assert that the statements made with respect to (7.5c) hold for L^1 and ϕ^1 , as well.

We shall now extend the property (1) to L^1 and K^1 (cf. (6.8), (6.8a)). Now in (7.11a) $K^{\delta_0, r}(x, y) \in H_{m-1}$; thus, by hypothesis, given any $\lambda = \lambda_0 = \alpha_0 + i\beta_0$ ($\beta_0 \neq 0$), there exists an operator $T_0^{\delta_0, r}$ (depending on λ_0 , independent of f) so that $\phi^{\delta_0, r}(x) = T_0^{\delta_0, r}(f(x)) \in L_2$ will be a solution of (7.11a) for all $f(x) \in L_2$ and so that

$$(7.12) \quad \int_a^b T_0^{\delta_0, r}(f_2(x)) f_1(x) dx = \int_a^b T_0^{\delta_0, r}(f_1(x)) f_2(x) dx \quad (\text{for } f_1, f_2 \in L_2).$$

The δ_0, r can be so chosen that

$$(7.12a) \quad T_0^{\delta_0, r}(f(x)) \rightarrow \phi^1(x) \quad (\text{in the weak sense}),$$

where $\phi^1(x) \in L_2$ constitutes a solution of (7.2) (for K^1, L^1, λ_0), and so that the relationship (7.12a) holds for all $f(x) \in L_2$. The function $\phi^1(x)$ of (7.12a) will then be related to an operation T_0^1 ,

$$(7.12b) \quad T_0^1(f(x)) = \phi^1(x) \quad (T_0^1 \text{ independent of } f),$$

defined for all $f(x) \in L_2$. By Theorem 1.4 and since $T_0^{b_0, r}(f(x)) \in L_2$ and (7.12a), (7.12b) hold, from (7.12) by passing to the limit it is inferred that

$$(7.12c) \quad \int_a^b T_0^1(f_2(x))f_1(x)dx = \int_a^b T_0^1(f_1(x))f_2(x)dx,$$

whenever $f_1, f_2 \in L$. The extension of (1) to the class H_m is immediate.

To establish the first part of (2) we assume (6.9) (for L^1) and, using (7.12c), repeat the steps (6.10)–(6.11) with reference to L^1 and K^1 . It remains to demonstrate that the number of linearly independent solutions, included in L_2 , of (7.2a) (for L^1, K^1) is the same for all non-real λ (under (6.9) for L^1). This is inferred with the aid of the operator T_0^1 of (7.12b), using arguments of the type given in (C, proof of Theorem V_4).

The property (3) (for L^1, K^1) is established as in the text in connection with (6.12)–(6.13a).

The statement (4) is extended to the class H_m on the basis of the inequality of (7.3), which has been already demonstrated for kernels included in H_m .

A consequence of this extension is that we are now able to assert that the result stated with regard to (7.6), (7.6a) holds for the class H_m , as well.

Similarly, it is seen that (5) will hold for K^1 and L^1 .

THEOREM 7.1. *The statements made, from (7.1) to (7.6a) and (5) (inclusive), will hold true, with respect to the equations (7.2), (7.2a), for all classes H_n (finite n).*

8. Some further results for classes H_n . The following formulas (cf. (8.1)–(8.4)), which were established in (C) for the class H_1 , will hold for all classes H_n , provided that we envisage only kernels with which one may “associate” (Definition 2.3) operators L and provided that (7.2a) ($\Gamma \neq 0$) has $\phi(y) = 0$ (almost everywhere) as the only solution included in L_2 .

One has

$$(8.1) \quad \int_a^b \left[\frac{\partial}{\partial x} \int_{\Delta_1} \mu d_\mu \Omega(x, y | \mu) \right] \left[\frac{\partial}{\partial x} \Delta_2 \Omega(x, z | \mu) \right] dx \\ = \int_a^b \left[\frac{\partial}{\partial x} \int_{\Delta_2} \mu d_\mu \Omega(x, z | \mu) \right] \left[\frac{\partial}{\partial x} \Delta_1 \Omega(x, y | \mu) \right] dx.*$$

* If the difference operator Δ corresponds to the interval (λ', λ'') , integration extended over Δ will be understood to be between the limits λ', λ'' .

When the intervals corresponding to Δ_1, Δ_2 are nonoverlapping,

$$(8.2) \quad \int_a^b \left[\frac{\partial}{\partial x} \Delta_1 \Omega(x, y | \lambda) \right] \left[\frac{\partial}{\partial x} \Delta_2 \Omega(x, z | \lambda) \right] dx = 0;$$

moreover,

$$(8.3) \quad L_z \left(\xi \left| \frac{\partial}{\partial x} \Delta \Omega(x, z | \lambda) \right. \right) \\ = L_z \left(\xi \left| \frac{\partial}{\partial x} \int_a^b \frac{\partial}{\partial y} \Omega(x, y | \lambda) \frac{\partial}{\partial y} \Omega(y, z | \lambda) dy \right. \right).$$

When the only solution of $L(\xi | \phi(x)) = 0$ ($\phi(x) \in L_2$) is zero,

$$(8.4) \quad \Delta \Omega(x, z | \lambda) = \int_a^b \frac{\partial}{\partial y} \Omega(x, y | \lambda) \frac{\partial}{\partial y} \Omega(y, z | \lambda) dy.$$

To demonstrate (8.1) one may proceed as follows. By (3.27) the equations

$$(8.5) \quad L_z(\xi | \phi_1(x)) - \lambda \int_a^b L_z(\xi | K(x, s)) \phi_1(s) ds \\ = L_z \left(\xi \left| \frac{\partial}{\partial x} \int_{\Delta_1} (\mu - \lambda) d_\mu \Omega(x, y | \mu) \right. \right), \\ L_z(\xi | \phi_2(x)) - \lambda \int_a^b L_z(\xi | K(x, s)) \phi_2(s) ds \\ = L_z \left(\xi \left| \frac{\partial}{\partial x} \int_{\Delta_2} (\mu - \lambda) d_\mu \Omega(x, z | \mu) \right. \right)$$

possess solutions

$$(8.5a) \quad \phi_1(x) = \frac{\partial}{\partial x} \int_{\Delta_1} \mu d_\mu \Omega(x, y | \mu), \quad \phi_2(x) = \frac{\partial}{\partial x} \int_{\Delta_2} \mu d_\mu \Omega(x, z | \mu),$$

respectively. Now, by Theorem 7.1 the result stated in connection with (6.8) and (6.8a) holds for all classes H_n ; thus, on writing

$$f_1(x) = \frac{\partial}{\partial x} \int_{\Delta_1} (\mu - \lambda) d_\mu \Omega(x, y | \mu), \quad f_2(x) = \frac{\partial}{\partial x} \int_{\Delta_2} (\mu - \lambda) d_\mu \Omega(x, z | \mu),$$

it is inferred that

$$(8.5b) \quad \int_a^b \left[\frac{\partial}{\partial x} \int_{\Delta_1} (\mu - \lambda) d_\mu \Omega(x, y | \mu) \right] \left[\frac{\partial}{\partial x} \int_{\Delta_2} \mu d_\mu \Omega(x, z | \mu) \right] dx \\ = \int_a^b \left[\frac{\partial}{\partial x} \int_{\Delta_2} (\mu - \lambda) d_\mu \Omega(x, z | \mu) \right] \left[\frac{\partial}{\partial x} \int_{\Delta_1} \mu d_\mu \Omega(x, y | \mu) \right] dx.$$

Finally, (8.1) is obtained if one takes $\lambda = 1 + i\beta$ ($\beta \neq 0$) and equates the imaginary parts of the two members in (8.5b).

In demonstrating the property (8.2) for any class H_n one may follow a method analogous to that indicated in (C) for the class H_1 . We shall not give the details.

The identity (8.3) may be established by induction by consecutive passages to the limit.

The property (8.4) is a consequence of (8.3).

The results (8.1)–(8.4), which have been verified for all classes H_n ($n = 1, 2, \dots$) are of interest in themselves as well as with a view to further developments for the case when operators L , of a more specialized character than required by Definition 2.3, are available.

9. Regarding reducible sets. In the sequel, throughout, we let β denote a number of class I or II.* Let E be a nondense closed set on (a, b) with a denumerable derivative E^1 . Then E will be denumerable and we may write

$$(9.1) \quad E = (I_1, I_2, \dots),$$

$$(9.1a) \quad E^1 = (I_1^1, I_2^1, \dots);$$

moreover, E will be reducible and the derivative of order β will be zero,

$$(9.1b) \quad E^\beta = 0,$$

for some β of class I or II (we take β as the least number so that (9.1b) holds).

In §§2–8 the case corresponding to β of class I has been already considered. This is the reason why our attention will now be confined to the case of β (in (9.1b)) of class II. Necessarily β will be not a limit number.†

We shall need the following result.

Let G_1, G_2, \dots be a simply infinite‡ sequence of closed sets, each containing the next and each having some points not in the next. Let

$$(9.2) \quad G = G_1 G_2 \dots \subset O,$$

where O is an open set. Then either $G_1 \subset O$ or there exists a number j so that

$$(9.2a) \quad G, \subset O \quad (v = j + 1, j + 2, \dots),$$

while

$$(9.2b) \quad G_j \not\subset O. §$$

* It is to be recalled that the numbers of class I are the ordinals $1, 2, \dots$. The numbers not of the first class, but obtainable by the use of the two Cantor generation principles, are of class II. As usual ω will denote the first number of class II.

† That is, there will be a number $\beta - 1$.

‡ An infinite sequence $q_1, q_2, \dots, q_n, \dots$ is simply infinite if $n < \omega$.

§ $\not\subset$ in (9.2b) signifies that G_j has points not in O .

Suppose the above is not true. Then every set G_ν ($\nu = 1, 2, \dots$) has a point b_ν exterior to O . The point b_1 will be in each of the sets G_1, G_2, \dots, G_{n_1} and will be not in G_{n_1+1} ; in fact, if this were not the case b_1 would be in G and, by (9.5), it would be in O . The point b_{n_1+1} (exterior to O), being in G_{n_1+1} , will be distinct from b_1 ; b_{n_1+1} will belong to the sets

$$G_{n_1+1}, G_{n_1+2}, \dots, G_{n_2} \quad [b_{n_1+1} \in G_{n_2+1}],$$

since otherwise one would have

$$b_{n_1+1} \in G_{n_1+1} G_{n_1+2} \dots = G \subset O,$$

which presents a contradiction. Thus, step by step we obtain an infinite sequence of points $\{b_{n_i+1}\}$ ($0 = n_0 < n_1 < \dots$), which are all distinct and are all exterior to O , with the point b_{n_i+1} belonging to the finite number of sets

$$(9.3) \quad G_{n_i+1}, G_{n_i+2}, \dots, G_{n_{i+1}},$$

and not belonging to $G_{n_{i+1}+1}$ (such a point is obtained for $i = 0, 1, \dots$). Let

$$(9.3a) \quad \{c_k\} \quad (k = 1, 2, \dots)$$

be a subsequence of $\{b_{n_i+1}\}$ such that

$$(9.3b) \quad \lim_k c_k = c \quad [c_k = b_{n_{i'_k}+1}; i' = i_k; i_1 < i_2 < \dots]$$

exists.

Now, the points c_1, c_2, \dots are all in $G_{n_{i'_k}+1}$ ($i' = i_1$); the latter set being closed, c will be in it. In general, the points c_k, c_{k+1}, \dots will be all in $G_{n_{i'_k}+1}$ ($i' = i_k$) and hence, this set being closed,

$$(9.3c) \quad c \in G_{n_{i'_k}+1} \quad (i' = i_k).$$

The relation (9.3c) is asserted for $i' = i_1 < i_2 < \dots$. Clearly c belongs to every set G_ν ($\nu = 1, 2, \dots$) and thus is a point of G ; hence included in O . The latter set being open there exists a closed interval Δ , containing c in the interior and contained in O . In Δ there will be some points c_k ; that is, some points b_ν ; this is contrary to the italicized statement preceding (9.3). Whence we deduce the truth of the statement in connection with (9.2)–(9.2b).

Conforming with the notation introduced in §2, a set E satisfying (9.1b) (as stated) will be said to belong to $R_{\beta-1}$ (Definition 2.1). By definition of β the set $E^{\beta-1}$ will have some points; in view of (9.1b) the number of these points will be finite. Thus

$$(9.4) \quad E^{\beta-1} = (I_1^{\beta-1}, I_2^{\beta-1}, \dots, I_k^{\beta-1}) = (s_1, s_2, \dots, s_k) \quad (\beta - 1 \text{ of class II}).$$

We may also write

$$(9.5) \quad E^\alpha = (I_1^\alpha, I_2^\alpha, \dots) \quad (\alpha < \beta - 1).$$

The sets E^α ($\alpha < \beta - 1$) will be all denumerably infinite.

We form a set $\Delta^1(\delta_1)$ of closed intervals

$$(9.6) \quad \Delta^1(\delta_1) = (s_\nu - \delta_1, s_\nu + \delta_1) \quad (\nu = 1, \dots, k; \delta_1 > 0; \text{cf. (9.4)})$$

in such a way that they have no points in common and that no end point of them is a point of E . Given ϵ (> 0), however small, such a construction can always be effected with

$$(9.6a) \quad 0 < \delta_1 < \epsilon.$$

This is established using the fact that E is nondense. In fact, without loss of generality one may assume (for the purposes of demonstration) that the s_ν are interior to (a, b) and take ϵ sufficiently small so that the intervals*

$$(9.6b) \quad (s_\nu - \epsilon, s_\nu + \epsilon) \quad (\nu = 1, \dots, k)$$

are without common points and are interior to (a, b) . In the interval $(s_1, s_1 + \epsilon)$ we find a subinterval $(s_1 + a_1^1, s_1 + b_1^1)$ void of points of E ; in the interval $(s_1 - b_1^1, s_1 - a_1^1)$ we then find a subinterval $(s_1 - b_1, s_1 - a_1)$ free of points of E . We now consider the interval $(s_2 + a_1, s_2 + b_1) [\subset (s_2, s_2 + \epsilon)]$; in it is found another interval without any points of E , say $(s_2 + a_1^1, s_2 + b_1^1)$. Turning our attention to $(s_2 - b_1^1, s_2 - a_1^1)$ we find a subinterval $(s_2 - b_2, s_2 - a_2)$ free of points of E . Clearly, the intervals

$$(s_\nu - b_2, s_\nu - a_2), \quad (s_\nu + a_2, s_\nu + b_2) \quad (\nu = 1, 2)$$

will be void of points of E . Continuing in this manner one finally obtains numbers a_k, b_k so that

$$(s_\nu - b_k, s_\nu - a_k) \subset (s_\nu - \epsilon, s_\nu), \quad (s_\nu + a_k, s_\nu + b_k) \subset (s_\nu, s_\nu + \epsilon) \quad (\nu = 1, 2, \dots, k),$$

and so that the intervals here displayed in the first members are free of points of E . Accordingly, if one takes

$$a_k < \delta_1 < b_k,$$

all of the conditions stated in connection with (9.6), (9.6a) will be satisfied. A choice of the δ_ν , according to the above scheme, will be implied throughout in the sequel.

Inasmuch as β is not a limit number, there exists a limit number $\gamma \leq \beta - 1$ so that there exists no limit number τ for which $\gamma < \tau < \beta$. The sets corresponding to $\beta - 1, \beta - 2, \dots, \gamma + 1$,

$$(9.7) \quad E^{\beta-1}, E^{\beta-2}, \dots, E^{\gamma+1}, E^\gamma$$

* Unless stated otherwise, all the intervals will be supposed to be closed.

can be covered in succession by the $\beta - \gamma$ sets

$$(9.7a) \quad \Delta^1(\delta_1), \Delta^2(\delta_2), \dots, \Delta^{\beta-\gamma-1}(\delta_{\beta-\gamma-1}), \Delta^{\beta-\gamma}(\delta_{\beta-\gamma}),$$

each set (9.7a) to consist of a finite number of closed intervals, *the totality of all intervals, involved in (9.7a), being without common points, no end point of any of these intervals being coincident with any point of E* . The consecutive sets (9.7a) are constructed following the procedure described in §2. Thus, $\Delta^2(\delta_2)$ will consist of the intervals

$$\Delta^2_\nu(\delta_2) = (s_\nu^{\beta-2} - \delta_2, s_\nu^{\beta-2} + \delta_2) \quad (\nu = 1, \dots, m(\delta_1)),$$

where the $s_\nu^{\beta-2}$ ($1 \leq \nu \leq m(\delta_1)$) are the points of $E^{\beta-2}$ exterior to the set $\Delta^1(\delta_1)$. The set $\Delta^3(\delta_3)$ will consist of the intervals

$$\Delta^3_\nu(\delta_3) = (s_\nu^{\beta-3} - \delta_3, s_\nu^{\beta-3} + \delta_3)$$

$[\nu = 1, \dots, m(\delta_1, \delta_2)$; the $s_\nu^{\beta-3}$ are points of $E^{\beta-3}$ exterior to the set $\Delta^1(\delta_1) + \Delta^2(\delta_2)$]. The set $\Delta^{\beta-\gamma-1}(\delta_{\beta-\gamma-1})$ will consist of the intervals

$$(9.7b) \quad \Delta^{\beta-\gamma-1}_\nu(\delta_{\beta-\gamma-1}) = (s_\nu^{\gamma+1} - \delta_{\beta-\gamma-1}, s_\nu^{\gamma+1} + \delta_{\beta-\gamma-1}) \\ [\nu = 1, \dots, m(\delta_1, \delta_2, \dots, \delta_{\beta-\gamma-2})];$$

where the $s_\nu^{\gamma+1}$ are points of $E^{\gamma+1}$ exterior to the set $\Delta^1(\delta_1) + \dots + \Delta^{\beta-\gamma-2}(\delta_{\beta-\gamma-2})$. Since $E^{\gamma+1}$ is the derivative of E^γ it is observed that the limiting points of E^γ are all interior to

$$(9.7c) \quad \Delta^1(\delta_1) + \Delta^2(\delta_2) + \dots + \Delta^{\beta-\gamma-1}(\delta_{\beta-\gamma-1});$$

only a finite number of the points of E^γ , say

$$(9.7d) \quad s_1^\gamma, \dots, s_{m^1}^\gamma \quad (m^1 = m(\delta_1, \delta_2, \dots, \delta_{\beta-\gamma-1})),$$

will be in the open set (a, b) minus the set of (9.7c). Hence the points (9.7d) can be covered by the set $\Delta^{\beta-\gamma}(\delta_{\beta-\gamma})$, consisting of intervals

$$(9.7e) \quad \Delta^{\beta-\gamma}_\nu(\delta_{\beta-\gamma}) = (s_\nu^\gamma - \delta_{\beta-\gamma}, s_\nu^\gamma + \delta_{\beta-\gamma}) \quad [\nu = 1, \dots, m^1 \text{ (cf. (9.7d))}].$$

On taking account of the italics subsequent to (9.7a) it is clear that *the points belonging to the sets (9.7) are all in the open set*

$$(9.7f) \quad \Gamma(\delta_1, \delta_2, \dots, \delta_{\beta-\gamma}),$$

obtained by taking the sum of all the intervals involved in (9.7a) and discarding the end points of these intervals.

Suppose now that the limit number γ , obtained above, is

$$(9.8) \quad \gamma = 1 \cdot \omega.$$

In view of the results (9.2)–(9.2b), on noting that $E^\omega \subset O_\omega$, where O_ω is the

open set (9.7f), it is inferred that either all the sets of the simply infinite sequence*

$$(9.8a) \quad E, E^1, \dots, E^n, \dots$$

are in O_ω or there exists a number j [$=j(\delta_1, \delta_2, \dots, \delta_{\beta-\gamma}) < \omega$] so that

$$(9.8b) \quad E^{j+\nu} \subset O_\omega \quad (\nu = 1, 2, \dots),$$

while

$$(9.8c) \quad E^j \not\subset O_\omega.$$

The derivative E^{j+1} of E^j being contained in O_ω , only a finite number of points of E^j , say

$$(9.8d) \quad s_1^j, \dots, s_{m^1}^j \quad [m^1 = m(\delta_1, \delta_2, \dots, \delta_{\beta-\gamma})],$$

will be exterior to O_ω , no point (9.8d) being coincident with any end point of the intervals constituting O_ω . The points (9.8d) can be covered by the set $\Delta^{\beta-\gamma+1}(\delta_{\beta-\gamma+1})$, consisting of the intervals

$$(9.8e) \quad \Delta^{\beta-\gamma+1}(\delta_{\beta-\gamma+1}) = (s_\nu^j - \delta_{\beta-\gamma+1}, s_\nu^j + \delta_{\beta-\gamma+1}) \\ [\nu = 1, \dots, m^1; m^1 \text{ from (9.8d)}]$$

in such a way that the intervals in (9.7a), (9.8e) have no points in common, while no point of E is an end point of these intervals. Following the procedure of §2 we find that the remaining sets

$$E^{j-1}, E^{j-2}, \dots, E^1, E^0 = E$$

are covered by the sets

$$(9.9) \quad \Delta^{\beta-\gamma+2}(\delta_{\beta-\gamma+2}), \Delta^{\beta-\gamma+3}(\delta_{\beta-\gamma+3}), \dots, \Delta^{\beta-\gamma+i}(\delta_{\beta-\gamma+i}), \Delta^{\beta-\gamma+i+1}(\delta_{\beta-\gamma+i+1}).$$

Each set (9.8e), (9.9) will consist of a finite number of intervals

$$(9.9a) \quad \Delta^{\beta-\gamma+i}(\delta_{\beta-\gamma+i}) = (s_\nu^{j+1-i} - \delta_{\beta-\gamma+i}, s_\nu^{j+1-i} + \delta_{\beta-\gamma+i}) \\ [\nu = 1, \dots, m^1; m^1 = m(\delta_1, \delta_2, \dots, \delta_{\beta-\gamma+i-1}); i = 1, 2, \dots, j+1],$$

where s_ν^j is in E^j ; moreover, the construction is so effected that the closed intervals (which are finite in number), involved in the sequence of sets

$$(9.10) \quad \Delta^1(\delta_1), \Delta^2(\delta_2), \dots, \Delta^{\beta-\gamma+i+1}(\delta_{\beta-\gamma+i+1}),$$

are all without common points and that no point of E is an end point of these intervals. It is to be noted that the choice of δ_ν (>0) ($1 < \nu \leq \beta-\gamma+j+1$) depends on that of $\delta_1, \delta_2, \dots, \delta_{\nu-1}$; however, the choice of $\delta_1, \delta_2, \dots, \delta_{\nu-1}$ once

* By definition $E^\omega = E^1 E^2 \dots E^n \dots$ ($n < \omega$).

made, one may take δ , arbitrarily small and thus one may let δ , approach zero through suitable values.

We thus obtained a type of a covering theorem for the set E in the case when the limit number γ involved in (9.7f) is of the form $1 \cdot \omega$.

Suppose now that a covering theorem of the above description holds for all sets E for which

$$(9.11) \quad \gamma = \eta\omega, \quad 1 \leq \eta < \alpha,$$

where α is a number of 1st or 2d class. We wish to establish such a theorem for $\eta = \alpha$.

CASE 1. α is not a limit number. In this case we make use of the fact that the theorem holds for $\alpha - 1$.

CASE 2. α is a limit number. In considering Case 1 it is noted that, by hypothesis, there exists a finite number of sets

$$(9.12) \quad \Delta^1(\delta_1), \Delta^2(\delta_2), \dots, \Delta^H(\delta_H),$$

each consisting of finite number of closed intervals; the totality of these intervals will be without common points; every point of E will be an interior point of one of these intervals; it may occur that some of the subscripts in (9.12) ($1 < \nu < H$) depend on $\delta_1, \delta_2, \dots, \delta_{\nu-1}$. The number β (cf. (9.1b)) will be, of course, of the form

$$(9.12a) \quad \beta = (\alpha - 1)\omega + q \quad (0 < q < \omega);$$

moreover, according to the hypothesis, a covering theorem of the stated type will hold for every β (α fixed) with $q = 1, 2, \dots$, where $q < \omega$. It is desired now to obtain such a result when β has a value

$$(9.12b) \quad \beta^* = \alpha\omega + p \quad (0 < p < \omega).$$

With $E^{\beta^*} = 0$, E^{β^*-1} will consist of a finite number of intervals which can be covered by a set $\Delta^{*1}(\delta_1^*)$ of intervals; in succession we construct sets

$$(9.13) \quad \Delta^{*1}(\delta_1^*), \Delta^{*2}(\delta_2^*), \dots, \Delta^{*p}(\delta_p^*),$$

analogous to the sets (9.7a), and with similar properties. In particular, the last set in (9.13) will consist of the intervals

$$(9.13a) \quad \Delta^{*p}(\delta_p^*) = (s_{\nu}^{\alpha\omega} - \delta_p^*, s_{\nu}^{\alpha\omega} + \delta_p^*) \quad [\nu = 1, \dots, m^1],$$

where $m^1 = m(\delta_1^*, \delta_2^*, \dots, \delta_{p-1}^*)$. The $s_{\nu}^{\alpha\omega}$ in (9.13a) will be points of $E^{\alpha\omega}$; all the other points of $E^{\alpha\omega}$ (an infinity of them) will be interior to the set

$$\Delta^{*1}(\delta_1^*) + \dots + \Delta^{*p-1}(\delta_{p-1}^*).$$

By definition

$$(9.14) \quad E^{\alpha\omega} = E^{(\alpha-1)\omega+1} E^{(\alpha-1)\omega+2} \dots E^{(\alpha-1)\omega+n} \dots \quad (n < \omega).$$

Also, in view of the preceding it is observed that

$$(9.14a) \quad E^{\alpha\omega} \subset O,$$

where O is the open set obtained by discarding in

$$\Delta^{*1}(\delta_1^*) + \dots + \Delta^{*p}(\delta_p^*)$$

the end points of the intervals involved. On taking note of the result (9.2)–(9.2b), as applied with $G_\nu = E^{(\alpha-1)\omega+\nu}$, it is accordingly concluded that for some finite q (> 0)

$$E^{(\alpha-1)\omega+\nu} \subset O \quad (\nu = q, q+1, \dots),$$

while

$$(9.14b) \quad E^{(\alpha-1)\omega+q-1} \not\subset O,$$

unless $E^{(\alpha-1)\omega+1} \subset O$. It is arranged so that no points of $E^{(\alpha-1)\omega+q-1}$ is a boundary point of O . Only a finite number of points of $E^{(\alpha-1)\omega+q-1}$, say

$$(9.14c) \quad s_\nu^{(\alpha-1)\omega+q-1} \quad [\nu = 1, 2, \dots, m^1; m^1 = m(\delta_1^*, \dots, \delta_p^*)],$$

will be exterior to (9.14b).[†] If we consider the closed set

$$(9.15) \quad G = E \cdot [(a, b) - O],$$

the consecutive derivatives of G will be

$$(9.15a) \quad G^\nu = E^\nu \cdot [(a, b) - O] \quad (\nu = 1, 2, \dots, (\alpha-1)\omega + q),$$

where $G^{(\alpha-1)\omega+q-1}$ consists of the points (9.14c) and, consequently,

$$(9.15b) \quad G^\beta = 0 \quad (\beta = (\alpha-1)\omega + q).$$

In view of the hypothesis made in conjunction with (9.12), applying the statement, just referred to, to the set G , we obtain a finite number of sets (9.12) covering G , as stated subsequently to (9.12).[‡]

Every point of G is interior to $(a, b) - O$; G being closed, the sets (9.12) (covering G) can be replaced by subsets

$$(9.16) \quad \Delta_1^1(\delta_1), \Delta_1^2(\delta_2), \dots, \Delta_1^H(\delta_H),$$

respectively, obtained by replacing the intervals involved in (9.12), whenever necessary, by suitable subintervals in such a manner that not only the properties (with respect to (9.16)) of (9.12) are maintained but also every closed interval involved in (9.16) is interior to $(a, b) - O$. The finite sequence of sets (cf. (9.13))

[†] That is, will be interior points of $(a, b) - O$.

[‡] In the aforesaid statement replace E by G of (9.15).

$$\Delta^{*1}(\delta_1^*), \Delta^{*2}(\delta_2^*), \dots, \Delta^{*p}(\delta_p^*), \Delta_1^1(\delta_1), \Delta_1^2(\delta_2), \dots, \Delta_1^H(\delta_H)$$

will have all the required covering properties with respect to the set E , for which $E^{\beta^*} = 0$ (cf. (9.12b)). In other words, if Case 1 (introduced subsequent to (9.11)) is on hand and if the theorem holds for $\alpha - 1$, then it will also hold for α (α from (9.11)).

We now consider Case 2, *when α in (9.11) is a limit number*. Thus, it is assumed that every set E with $E^{\beta} = 0$, where

$$(9.17) \quad \beta = \eta\omega + q \quad (1 \leq \eta < \alpha; 1 \leq q < \omega), \dagger$$

can be "covered" by a finite number of sets

$$(9.17a) \quad \Delta^1(\delta_1), \Delta^2(\delta_2), \dots, \Delta^H(\delta_H),$$

each consisting of a finite number of intervals. Let \bar{E} be a set, with $\bar{E}^{\beta^*} = 0$, where

$$(9.17b) \quad \beta^* = \alpha\omega + p \quad (0 < p < \omega).$$

It is observed that $\bar{E}^{\alpha\omega}$ could be considered as the set common to all the sets

$$(9.17c) \quad \bar{E}^{\eta\omega} \quad (1 \leq \eta < \alpha).$$

We again form a finite number of sets (cf. (9.13)-(9.14a))

$$(9.17d) \quad \Delta^{*1}(\delta_1^*), \dots, \Delta^{*p}(\delta_p^*),$$

each consisting of a finite number of intervals; the last set displayed will be of the form (9.13a), where the $s, \alpha\omega$ (finite in number) are points of $\bar{E}^{\alpha\omega}$, the other points of $\bar{E}^{\alpha\omega}$ being interior points of $\Delta^{*1}(\delta_1^*) + \dots + \Delta^{*p-1}(\delta_{p-1}^*)$. The sets (9.17d) "cover" the sets $\bar{E}^{\beta^*-1}, \bar{E}^{\beta^*-2}, \dots, \bar{E}^{\alpha\omega}$. We again have

$$(9.18) \quad \bar{E}^{\alpha\omega} \subset O,$$

where O is the open set, obtained by taking the sum of the sets (9.17d) and discarding the end points of the intervals involved.

The sequence of sets

$$(9.19) \quad \bar{E}^{1\omega}, \bar{E}^{2\omega}, \dots, \bar{E}^{\eta\omega}, \dots \quad (\eta < \alpha),$$

even though denumerable, may be not a simply infinite sequence. It is not difficult to see that the set $\bar{E}^{\alpha\omega}$, which is the product of the sets (9.19), could also be considered as the product of the sets of the following *simply infinite* sequence:

† It is understood that $E^{\beta-1}$ has some points. In (9.17a) the last set contains just a finite number of points of E , all the other points of E being in the other sets of (9.17a).

$$(9.19a) \quad \bar{E}_{\eta_1\omega}, \bar{E}_{\eta_2\omega}, \dots \quad (\eta_1 < \eta_2 < \dots; \eta_r < \alpha),$$

provided the η_r are suitably chosen.†

With (9.18) and the above in view, the theorem (9.2)–(9.2b) can be applied with $G = \bar{E}^{\alpha\omega}$ and $G_r = \bar{E}_{\eta_r\omega}$. Thus there exists a finite number j so that

$$(9.20) \quad \bar{E}_{\eta_r\omega} \subset O \quad (\nu = j+1, j+2, \dots),$$

while

$$(9.20a) \quad \bar{E}_{\eta_j\omega} \not\subset O,$$

unless $\bar{E}_{\eta_j\omega} \subset O$. No points of $\bar{E}_{\eta_j\omega}$ will be coincident with any of the boundary points of O . Form the set $E_1 = \bar{E}_{\eta_j\omega}[(a, b) - O]$ and write

$$(9.21) \quad E = \bar{E}[(a, b) - O];$$

then

$$(9.21a) \quad E^\gamma = \bar{E}^\gamma[(a, b) - O] \quad (\gamma = 1, 2, \dots),$$

$$(9.21b) \quad E_{\eta_j\omega} = E_1.$$

Now, by (9.20)

$$(9.21c) \quad E_{\eta_{j+1}\omega} = 0.$$

Thus, for some $\beta = \eta_j\omega + \tau \leq \eta_{j+1}\omega$ we have

$$(9.21d) \quad E^\beta = E_1^\tau = 0.$$

Since $\eta_{j+1} < \alpha$ one has $\beta < \alpha\omega$; hence in (9.21d) β is of the form (9.17). Let (9.17a) constitute the set “covering” E . By taking suitable subintervals of the intervals constituting the sets (9.17a) we correspondingly obtain other “covering” sets:

$$(9.22) \quad \Delta_1^1(\delta_1), \Delta_1^2(\delta_2), \dots, \Delta_1^H(\delta_H), \ddagger$$

having the same properties as (9.17a) but which at the same time are interior to $(a, b) - O$. The reasoning in this connection is the same as that previously made in connection with (9.16). Adjoining the sequences (9.22), (9.17d), and

$$(9.23) \quad \Delta^{*1}(\delta_1^*), \Delta^{*2}(\delta_2^*), \dots, \Delta^{*p}(\delta_p^*), \Delta_1^1(\delta_1), \dots, \Delta_1^H(\delta_H),$$

we obtain a finite number of sets, each consisting of a finite number of closed intervals, the totality of these intervals possessing no common points, every point of \bar{E} being an interior point of one of the intervals involved; the se-

† Thus, for instance, if $\alpha = \omega^3$, one may take $\eta_r = r\omega^2$. With α a limit number, $\eta_r (< \alpha)$ must be chosen so that, given any $\gamma < \alpha$, one may find a value r ($r < \omega$) for which $\gamma < \eta_r < \alpha$.

‡ Here the δ_r may be different from the original ones.

quence (9.23) will "cover" the set \bar{E} in the sense previously attributed to the word "cover."

This completes the transfinite induction.

THEOREM 9.1. *Let E be a reducible set as stated at the beginning of this section (cf. (9.1)–(9.1b)). The sets*

$$E, E^1, E^2, \dots, E^{\beta-1} \quad (\beta \text{ of 1st or 2d class})^\dagger$$

can always be "covered," in the sense indicated above by a finite number of sets

$$(9.24) \quad \Delta^1(\delta_1), \Delta^2(\delta_2), \dots, \Delta^q(\delta_q) \quad (\delta_1 > 0, \dots, \delta_q > 0),$$

the set $\Delta^r(\delta_r)$ consisting of a finite number of intervals of length $2\delta_r$, no point of E being an end point of any of the intervals involved.

NOTE. It is observed that in (9.24) the choice of a particular δ_r ($r > 1$) depends on that of $\delta_1, \dots, \delta_{r-1}$; moreover, having chosen $\delta_1, \dots, \delta_{r-1}$, we may take the number δ_r arbitrarily small. It is also to be noted that the number q in (9.24) may depend on the choice of $\delta_1, \delta_2, \dots$.

10. Kernels of transfinite rank. Such kernels will be introduced by means of the following definition.

DEFINITION 10.1. *Let E be a closed reducible set on (a, b) , as described in the beginning of §9, with $E^\beta = 0$ (β a non-limit number of class II) and $E^{\beta-1}$ consisting of some points (necessarily finite in number). Let $\Delta^r(\delta_r)$ [$r = 1, 2, \dots, q$ ($< \omega$)] be the corresponding covering sets, referred to in Theorem 9.1. A kernel $K(x, y)$ will be said to belong to the class H_β if, for all "admissible" values δ_r (> 0 ; $r = 1, \dots, q$),*

$$(10.1) \quad K^{\delta_1, \delta_2, \dots, \delta_q}(x, y) \in L_2 \quad (\text{in } x, y; \text{ for } a \leq x, y \leq b).$$

Here

$$(10.2) \quad K^{\delta_1, \delta_2, \dots, \delta_q}(x, y) = 0 \quad [x \text{ in } \sum_{r=1}^q \Delta^r(\delta_r); a \leq y \leq b \text{ (or } a \leq y < x)];$$

$$(10.2a) \quad K^{\delta_1, \delta_2, \dots, \delta_q}(x, y) = 0 \quad [y \text{ in } \sum_{r=1}^q \Delta^r(\delta_r); a \leq x \leq b \text{ (or } a \leq x < y)];$$

$$(10.2b) \quad K^{\delta_1, \delta_2, \dots, \delta_q}(x, y) = K(x, y) \quad [\text{at all other points of } a \leq x, y \leq b].$$

We get in succession

$$(10.3) \quad \begin{aligned} \lim_{\delta_q} K^{\delta_1, \dots, \delta_q}(x, y) &= K^{\delta_1, \dots, \delta_{q-1}}(x, y), \\ \lim_{\delta_{q-1}} K^{\delta_1, \dots, \delta_{q-1}}(x, y) &= K^{\delta_1, \dots, \delta_{q-2}}(x, y), \dots, \\ \lim_{\delta_2} K^{\delta_1, \delta_2}(x, y) &= K^{\delta_1}(x, y), \quad \lim_{\delta_1} K^{\delta_1}(x, y) = K(x, y). \end{aligned}$$

Thus $K(x, y)$ is a q -fold repeated limit of the function of the first member in

[†] As noted before, necessarily β is not a limit number.

|| That is, if h represents a point of H , the corresponding point of $F_i(H)$ will be represented by $1/(i+1) + (1/i - 1/(i+1))h$.

$$(10.6) \quad E^n = O + F_n(O) + F_{n+1}(A_1) + F_{n+2}(A_2) + \dots$$

for $n=1, 2, \dots (n < \omega)$. It is noted that E^n is in the closed interval $(0, 1/(n+1))$, the point $1/(n+1)$ being an isolated point of E^n .

To cover E , as defined by (10.5a), by a finite number of sets of intervals, following the scheme which we established previously, we proceed as follows. The set $\Delta^1(\delta_1)$ will consist of the single interval

$$(10.6a) \quad \Delta^1(\delta_1) = (0, \delta_1) \quad (0 < \delta_1 < 1/2),$$

where δ_1 is not coincident with any of the points of E . The set (10.6a) "covers" E^n . Corresponding to δ_1 there exists a number $j=j(\delta_1)$ [$j(\delta_1) \rightarrow \infty$ with $1/\delta_1$], such that every point of each of the sets

$$(10.6b) \quad E^{j+1}, E^{j+2}, \dots$$

is interior to* $(0, \delta_1)$, while the set E^j has points not in $\Delta^1(\delta_1)$; that is, there are points of E^j in the interval $(\delta_1 < x < 1)$. In view of the statement subsequent to (10.6) one clearly has

$$\frac{1}{j+2} < \delta_1 < \frac{1}{j+1};$$

that is,

$$(10.6c) \quad \frac{1}{\delta_1} - 2 < j < \frac{1}{\delta_1} - 1. \dagger$$

The points of E^j exterior to $\Delta^1(\delta_1)$ will be $F_j(O)$, that is $1/(j+1)$, and those points of $F_{j+1}(A_1)$ which are to the right of δ_1 . The points of A_1 being $0, 1/2, 1/3, \dots$, those of $F_{j+1}(A_1)$ will be

$$(10.6d) \quad \frac{1}{j+2} + l_{j+1}\nu \quad (l_{j+1} = 1/(j+1) - 1/(j+2)),$$

where $\nu=0, 1/2, 1/3, \dots$. Accordingly it is concluded that the points of E^j , exterior to $\Delta^1(\delta_1)$, will consist either of the single point $1/(j+1)$, or of the m^1 (>1) points

$$(10.6e) \quad s_1^j = \frac{1}{j+1}, \quad s_2^j = \frac{1}{j+2} + \frac{1}{2} l_{j+1},$$

$$s_3^j = \frac{1}{j+2} + \frac{1}{3} l_{j+1}, \dots, \quad s_{m^1}^j = \frac{1}{j+2} + \frac{1}{m^1} l_{j+1},$$

where $m^1 = m(\delta_1)$ and

* According to previously made conventions, "interior to" here means "in the interval $a \leq x < \delta_1$."

† Given an "admissible" δ_1 this defines the integer j uniquely.

$$(10.6f) \quad \frac{l_{j+1}}{\delta_1 - (j+2)^{-1}} - 1 < m^1 < \frac{l_{j+1}}{\delta_1 - (j+2)^{-1}}.$$

In succession "covering" sets $\Delta^r(\delta_r)$ ($r=2, 3, \dots, j+2$) are obtained, with $\Delta^r(\delta_r)$ consisting of intervals

$$(10.7) \quad \Delta^r(\delta_r) = (s_\nu^{j+2-r} - \delta_r, s_\nu^{j+2-r} + \delta_r) \\ [\nu = 1, 2, \dots, m^1; m^1 = m(\delta_1, \delta_2, \dots, \delta_{r-1})],$$

so that E^j is interior to $\Delta^1(\delta_1) + \Delta^2(\delta_2)$ (a finite number of points of E^j in $\Delta^2(\delta_2)$), E^{j-1} is interior to $\Delta^1(\delta_1) + \dots + \Delta^3(\delta_3)$ (a finite number of points of E^{j-1} in $\Delta^3(\delta_3)$) and so on, with $E = E^0$ interior to the set

$$(10.7a) \quad \Delta^1(\delta_1) + \dots + \Delta^{j+2}(\delta_{j+2}),$$

only a finite number of points of E lying in $\Delta^{j+2}(\delta_{j+2})$.*

Let

$$(10.8) \quad K(x, y) = g(x) \quad (\text{for } 0 \leq y < x),$$

$$(10.8a) \quad K(x, y) = g(y) \quad (\text{for } 0 \leq x < y),$$

where $g(x)$ is defined as follows. The set E of (10.5a) being closed nondense, the complementary set $(0, 1) - E$ will consist of a denumerable infinity of open (except at 1), nonoverlapping intervals

$$(10.8b) \quad (a_i, b_i) \quad (0 \leq a_i < b_i \leq 1; i = 1, 2, \dots),$$

some of them being adjacent. We let $g(x) = 0$ for $3/4 \leq x \leq 1$ and in the other intervals (10.8b) we take

$$(10.8c) \quad g(x) = 0 \quad (\text{for } a_i < x < (a_i + b_i)/2), \\ g^2(x) = \frac{c_i}{b_i^2 - x^2} \quad (\text{for } (a_i + b_i)/2 \leq x < b_i; c_i > 0).^\dagger$$

In accordance with Definition 10.1 related to $K(x, y)$ will be the kernel

$$(10.9) \quad K^{\delta_1, \dots, \delta_{j+2}}(x, y) = g^{\delta_1, \dots, \delta_{j+1}}(x) \quad (0 \leq y < x),$$

$$(10.9a) \quad \zeta^{\delta_1, \dots, \delta_{j+2}}(x, y) = g^{\delta_1, \dots, \delta_{j+1}}(y) \quad (0 \leq x < y),$$

where

$$(10.9b) \quad g^{\delta_1, \dots, \delta_{j+2}}(x) = 0 \quad [\text{for } x \text{ in } \sum_{\nu=1}^{j+2} \Delta^\nu(\delta_\nu)], \\ g^{\delta_1, \dots, \delta_{j+2}}(x) = g(x) \quad [\text{for } x \text{ in } (0, 1) - \sum_{\nu=1}^{j+2} \Delta^\nu(\delta_\nu)].$$

* The numbers m^1 involved in (10.7) can be specified by inequalities in terms of $\delta_1, \dots, \delta_{r-1}$ in succession, making use of the fact that the δ_ν are chosen so as to secure "covering" in our sense.

† The definition for x equal to a number corresponding to a point of E is immaterial.

Now the G_i in (10.8c) are points in E and they are interior to the finite number of nonadjacent nonoverlapping intervals constituting (10.7a). The set

$$(10.9c) \quad (0, 1) - \sum_1^{j+2} \Delta^v(\delta_v)$$

will consist of a finite number of open intervals Γ . Any particular interval Γ will be, together with its end points, interior to some interval (a_i, b_i) . In view of (10.8c) and (10.9b)

$$|g^{\delta_1, \dots, \delta_{j+2}}(x)|^2 \leq \frac{|c_i|}{|b_i^2 - x^2|} \quad (x \text{ in interval } \Gamma).$$

With b_i exterior to the given interval Γ , $|g^{\delta_1, \dots, \delta_{j+2}}(x)|$ will be uniformly bounded in this interval. Hence this function will be uniformly bounded in the total set of intervals constituting Γ , inasmuch as the number of these intervals is finite. On taking account of the first relation (10.9b) it is finally concluded that

$$(10.9d) \quad |g^{\delta_1, \dots, \delta_{j+2}}(x)| \leq \beta(\delta_1, \dots, \delta_{j+2}) \quad (0 \leq x \leq 1),$$

where the second member is independent of x and is finite whenever $\delta_1, \dots, \delta_{j+2}$ have positive "admissible" values. In consequence of (10.9), (10.9a), and (10.9d) $K^{\delta_1, \dots, \delta_{j+2}}(x, y) \in L_2$ (in x, y ; for $0 \leq x, y \leq 1$). Thus, (10.8), (10.8a) furnishes an example of a kernel $K(x, y) \in H_{\omega+1}$. Now, we recall that convergence of

$$(10.10) \quad \int_0^1 K^2(x, y) dy \quad (\text{almost all } x \text{ on } (0, 1))$$

would imply that $K(x, y)$ is essentially of one of the classes of kernels considered by Carleman. *This, however, is not the case.* In fact, the integral (10.10), if convergent, could be written as

$$(10.10a) \quad xg^2(x) + \int_x^1 g^2(y) dy.$$

For every $0 \leq x < 3/4$ the interval $(x, 1)$ would contain at least one point b_i (cf. (10.8c)). The presence of such an infinite discontinuity implies that there exists no integral (10.10a).

On determining in succession the limits

$$(10.11) \quad \begin{aligned} \lim_{\delta_{j+2}} K^{\delta_1, \dots, \delta_{j+2}}(x, y) &= K^{\delta_1, \dots, \delta_{j+1}}(x, y), \\ \lim_{\delta_{j+1}} K^{\delta_1, \dots, \delta_{j+1}}(x, y) &= K^{\delta_1, \dots, \delta_j}(x, y), \dots, \\ \lim_{\delta_2} K^{\delta_1, \delta_2} &= K^{\delta_1}(x, y), \quad \lim_{\delta_1} K^{\delta_1}(x, y) = K(x, y), \end{aligned}$$

it is observed that

$$(10.11a) \quad \begin{aligned} K^{\delta_1, \dots, \delta_{j+1}}(x, y) &\in H_1, & K^{\delta_1, \dots, \delta_j}(x, y) &\in H_2, \dots, \\ K^{\delta_1, \delta_2}(x, y) &\in H_j, & K^{\delta_1}(x, y) &\in H_{j+1}. \end{aligned}$$

The latter relationship above implies that $K(x, y)$ is a limit of a simply infinite sequence of kernels each of which belongs to a class H_ν ($\nu < \omega$). It is of interest to assure that

$$(10.11b) \quad \int_0^1 |K^{\delta_1}(x, y)|^2 dy \quad (\text{all "admissible" } \delta_1 > 0)$$

diverges, since in the contrary case $K^{\delta_1}(x, y)$ would be essentially of Carleman's type and $K(x, y)$, itself, would be of rank two. Now

$$K^{\delta_1}(x, y) = \begin{cases} 0 & (0 \leq x \leq \delta_1; 0 \leq y < x), \\ 0 & (0 \leq y \leq \delta_1; 0 \leq x < y), \end{cases}$$

and $K^{\delta_1}(x, y) = K(x, y)$ at the other points of $0 \leq x, y \leq 1$.^{*} The integral (10.11b), if it exists, is of the form (10.10a), where $g(x)$ is replaced by $g^{\delta_1}(x)$,

$$g^{\delta_1}(x) = \begin{cases} 0 & (0 \leq x \leq \delta_1), \\ g(x) & (\delta_1 < x \leq 1). \end{cases}$$

Thus, the integral (10.11a) will exist if and only if

$$(10.11c) \quad \int_x^1 |g^{\delta_1}(y)|^2 dy$$

exists. The expression last displayed is identical with

$$\int_{\delta_1}^1 g^2(y) dy \quad (\text{for } 0 \leq x \leq \delta_1),$$

which diverges for $\delta_1 < 3/4$ (cf. the statement subsequent to (10.10a)). When $\delta_1 < x < 1$

$$\int_x^1 |g^{\delta_1}(y)|^2 dy = \int_x^1 g^2(y) dy;$$

this diverges for reasons previously given with reference to (10.10a). Hence it is inferred that (10.11b) diverges, as stated. It remains to make certain that $K(x, y)$ does not belong to a class H_n , where $n < \omega$. In fact, suppose $K(x, y) \in H_n$. Then by Definition 2.2† there exists a set $\bar{E} = \bar{E}^0$, with

^{*} In accordance with certain previous remarks, the values of $K(x, y)$ on lines $x = \text{number}$ corresponding to a point of E are immaterial.

† To conform with the present notation, δ_ν, Δ^ν , in Definition 2.2, are replaced by $\delta_{\nu+1}, \Delta^{\nu+1}$, respectively.

$$\bar{E}^{n-1} = (\bar{s}_1, \dots, \bar{s}_k),$$

and sets $\Delta_1^r(\bar{\delta}_r)$ ($r=1, \dots, n$) of intervals such that the points of \bar{E}^{n-i} are interior to $\Delta_1^1(\bar{\delta}_1) + \dots + \Delta_1^i(\bar{\delta}_i)$ ($i=1, \dots, n$), the only points of \bar{E}^{n-i} in $\Delta_1^i(\bar{\delta}_i)$ being the centers

$$\bar{s}_\nu^{n-i} \quad [\nu = 1, \dots, m^1; m^1 = m(\bar{\delta}_1, \bar{\delta}_2, \dots, \bar{\delta}_{i-1})]$$

of the constituent intervals of $\Delta_1^i(\bar{\delta}_i)$. Moreover,

$$(10.12) \quad K^{\bar{\delta}_1, \dots, \bar{\delta}_n}(x, y) \in L_2 \quad (\text{in } x, y);$$

here the first member equals $K(x, y)$, except for x in $G = \Delta_1^1(\bar{\delta}_1) + \dots + \Delta_1^n(\bar{\delta}_n)$ ($0 \leq y < x$) and also for y in G ($0 \leq x < y$), where the first member of (10.12) is zero. The complement of G could not contain an interval $(0, h)$ ($h > 0$), in fact, if it did one would have

$$\int_0^h \int_0^h |K^{\bar{\delta}_1, \dots, \bar{\delta}_n}(x, y)|^2 dx dy = \int_0^h \int_0^h K^2(x, y) dx dy.$$

The integrand last displayed has an infinity of infinite discontinuities within the field of integration, each of which, alone, would suffice to secure divergence of the integral in question. Thus (10.12) does not hold for all "admissible" sufficiently small $\bar{\delta}_r$, unless G contains an interval $(0, h)$ ($h > 0$) for every "admissible" choice of $\bar{\delta}_1, \dots, \bar{\delta}_n$.^{*} Hence the point 0 must be the center (that is, end point, in this case) of an interval of G . Suppose this point is the center of an interval Γ_1 of the set $\Delta_1^r(\bar{\delta}_r)$. We recall that having made an "admissible" choice of $\bar{\delta}_1, \bar{\delta}_2, \dots, \bar{\delta}_{r-1}$, the choice of $\bar{\delta}_r, \dots, \bar{\delta}_n$ depends on that of $\bar{\delta}_1, \dots, \bar{\delta}_{r-1}$; however, $\bar{\delta}_r$ may be taken arbitrarily small. Thus, with $\bar{\delta}_r$ suitably small, there will exist an interval Γ_2 , adjacent to and nonoverlapping with Γ_1 , which will be in the complement of G and in which $K(x, y)$ will have infinite discontinuities (cf. (10.8)–(10.8c)); on the other hand,

$$K^{\bar{\delta}_1, \dots, \bar{\delta}_n}(x, y) = K(x, y)$$

for x in Γ_2 and $0 \leq y \leq 1$ and also for y in Γ_2 and $0 \leq x \leq 1$. The presence of the above discontinuities implies that (10.12) does not hold, as stated. Thus, *our kernel $K(x, y)$, as given by (10.8)–(10.8c), is of the transfinite class $H_{\omega+1}$ and does not belong to any class of index less than $\omega+1$.*

Following the procedure indicated from (10.8) to the italicized statement above, obvious generalizations can be made regarding existence and construction of kernels of various transfinite classes.

11. Results for classes H_β . Let $K(x, y) \in H_\beta$ where $\beta = \omega + p$ ($0 < p < \omega$).

^{*} h may depend on $\bar{\delta}_1, \dots, \bar{\delta}_n$.

For the corresponding set E we shall have

$$(11.1) \quad E^{\omega+p-1} = (s_\nu) \quad (\nu = 1, \dots, k).$$

Sets $\Delta^r(\delta_\nu)$ each consisting of a finite number of intervals are constructed so that

$$(11.1a) \quad \begin{aligned} E^{\omega+p-1} &\subset \Delta^1(\delta_1), & E^{\omega+p-2} &\subset \Delta^1(\delta_1) + \Delta^2(\delta_2), \dots, \\ E^{\omega+1} &\subset \Delta^1(\delta_1) + \dots + \Delta^{p-1}(\delta_{p-1}), & E^\omega &\subset \Delta^1(\delta_1) + \dots + \Delta^p(\delta_p) = \Gamma, \end{aligned}$$

where $G \subset H$ signifies that every point of G is an interior point of H . In (11.1a) the only points $E^{\omega+p-i}$ which are in $\Delta^i(\delta_i)$ are the centers

$$(11.1b) \quad s_\nu^{\omega+p-i} \quad [\nu = 1, \dots, m^1; m^1 = m(\delta_1, \dots, \delta_{i-1})]^*$$

of the constituent intervals of $\Delta^i(\delta_i)$; this assertion is made for $i = 1, \dots, p$. Furthermore, for some $j = j(\delta_1, \dots, \delta_p)$

$$E^{j+\nu} \subset \Gamma \quad [\nu = 1, 2, \dots; \text{cf. (11.1a)}],$$

while

$$(11.1c) \quad E^j \not\subset \Gamma,$$

the points of E^j , not in Γ , being finite in number,

$$s_1^j, \dots, s_{m^1}^j \quad (m^1 = m(\delta_1, \dots, \delta_p)).$$

Further sets $\Delta^r(\delta_\nu)$ ($\nu = p+1, p+2, \dots, p+j+1$) of intervals are constructed so that

$$(11.2) \quad \begin{aligned} E^j &\subset \Gamma + \Delta^{p+1}(\delta_{p+1}), & E^{j-1} &\subset \Gamma + \Delta^{p+1}(\delta_{p+1}) + \Delta^{p+2}(\delta_{p+2}), \dots, \\ E^1 &\subset \Gamma + \Delta^{p+1}(\delta_{p+1}) + \dots + \Delta^{p+j}(\delta_{p+j}), \\ E^0 &= E \subset \Gamma + \Delta^{p+1}(\delta_{p+1}) + \dots + \Delta^{p+j+1}(\delta_{p+j+1}), \end{aligned}$$

the symbol \subset having the same meaning as in (11.1a), the only points of E^{j-r} in $\Delta^{p+r+1}(\delta_{p+r+1})$ being the centers

$$(11.2a) \quad s_1^{j-r}, s_2^{j-r}, \dots, s_{m^1}^{j-r} \quad [m^1 = m(\delta_1, \dots, \delta_{p+r})]$$

of the intervals constituting $\Delta^{p+r+1}(\delta_{p+r+1})$ ($r = 0, 1, \dots, j$).

According to Definition (10.1), associated with $K(x, y)$ will be the function

$$(11.3) \quad K^{\delta_1, \delta_2, \dots, \delta_q}(x, y) \quad (q = p + j + 1),$$

satisfying the conditions of that Definition. In succession we define the limits

* For $i = 1$, $m^1 = k$.

$$\begin{aligned}
 (11.3a) \quad & \lim_{\delta_{p+j+1}} K^{\delta_1, \dots, \delta_{p+j+1}}(x, y) = K^{\delta_1, \dots, \delta_{p+j}}(x, y), \\
 & \lim_{\delta_{p+j}} K^{\delta_1, \dots, \delta_{p+j}}(x, y) = K^{\delta_1, \dots, \delta_{p+j-1}}(x, y), \dots, \\
 & \lim_{\delta_{p+1}} K^{\delta_1, \dots, \delta_{p+1}}(x, y) = K^{\delta_1, \dots, \delta_p}(x, y),
 \end{aligned}$$

where, in particular,

$$\begin{aligned}
 (11.3b) \quad & K^{\delta_1, \dots, \delta_p}(x, y) = 0 \quad [\text{for } x \text{ in } \Gamma; 0 \leq y \leq 1], \\
 & = 0 \quad [\text{for } y \text{ in } \Gamma; 0 \leq x \leq 1; \text{ cf. (11.1a)}],
 \end{aligned}$$

and $K^{\delta_1, \dots, \delta_p}(x, y) = K(x, y)$ at the other points of $0 \leq x, y \leq 1$. It is essential to note that the kernels involved in (11.3a) are all of finite ranks. The next limiting process

$$(11.3c) \quad \lim_{\delta_p} K^{\delta_1, \dots, \delta_p}(x, y) = K^{\delta_1, \dots, \delta_{p-1}}(x, y)$$

is essentially distinct from those of (11.3a); it yields a kernel which may be actually of transfinite rank.* Further limiting processes will yield

$$\begin{aligned}
 (11.3d) \quad & \lim_{\delta_{p-1}} K^{\delta_1, \dots, \delta_{p-1}}(x, y) = K^{\delta_1, \dots, \delta_{p-2}}(x, y), \dots, \\
 & \lim_{\delta_2} K^{\delta_1, \delta_2}(x, y) = K^{\delta_1}(x, y), \quad \lim_{\delta_1} K^{\delta_1}(x, y) = \nabla^0(x, y) = K(x, y).
 \end{aligned}$$

Clearly

$$(11.4) \quad |K^{\delta_1, \dots, \delta_{p+1}}(x, y)| \leq |K^{\delta_1, \dots, \delta_p}(x, y)| \quad (p = 0, 1, \dots, p+j).$$

Corresponding to the last member of (11.3a) we form the equations

$$(11.5) \quad \phi^{\delta_1, \dots, \delta_p}(x) - \lambda \int_a^b K^{\delta_1, \dots, \delta_p}(x, y) \phi^{\delta_1, \dots, \delta_p}(y) dy = f(x) \quad (f(x) \in L_2),$$

$$(11.6) \quad \phi^{\delta_1, \dots, \delta_p}(x) - \lambda \int_a^b K^{\delta_1, \dots, \delta_p}(x, y) \phi^{\delta_1, \dots, \delta_p}(y) dy = 0.$$

Inasmuch as $K^{\delta_1, \dots, \delta_p}(x, y)$ ($\delta_1 > 0, \dots, \delta_p > 0$) is of some class H_n ($n < \omega$) the results of Theorem 4.1 will be applicable to the equations (11.5), (11.6).

Thus, corresponding to $K^{\delta_1, \dots, \delta_p}(x, y)$ there exists a function $\Omega^{\delta_1, \dots, \delta_p}(x, y)$ such that

$$(11.7) \quad \text{Var. } \Omega^{\delta_1, \dots, \delta_p}(x, y | \lambda) \leq [(x-a)(y-a)]^{1/2}, \quad \Omega^{\delta_1, \dots, \delta_p}(x, y | 0) = 0,$$

$$\begin{aligned}
 (11.7a) \quad & |\Omega^{\delta_1, \dots, \delta_p}(x^1, y^1 | \lambda) - \Omega^{\delta_1, \dots, \delta_p}(x, y | \lambda)| \\
 & \leq (b-a)^{1/2} (|y^1 - y|^{1/2} + |x^1 - x|^{1/2});
 \end{aligned}$$

* That is, in some cases $K^{\delta_1, \dots, \delta_{p-1}}(x, y)$ will belong to $H_{\omega+1}$ and will be not of class H_ν ($\nu < \omega$).

this function may be discontinuous in λ for a denumerable infinity of real values λ . In view of (11.7), (11.7a), application of the "Compactness Theorem" (§1), with respect to δ_p , leads to the conclusion that the limit

$$(11.8) \quad \lim_{\delta_{p,r}} \Omega^{\delta_1, \dots, \delta_p}(x, y | \lambda) = \Omega^{\delta_1, \dots, \delta_{p-1}}(x, y | \lambda) \quad [\text{suitable } \delta_{p,r}; \lim \delta_{p,r} = 0]$$

exists and satisfies

$$(11.8a) \quad \text{Var. } \Omega^{\delta_1, \dots, \delta_{p-1}}(x, y | \lambda) \leq [(x-a)(y-a)]^{1/2}, \Omega^{\delta_1, \dots, \delta_{p-1}}(x, y | 0) = 0, \\ |\Omega^{\delta_1, \dots, \delta_{p-1}}(x^1, y^1 | \lambda) - \Omega^{\delta_1, \dots, \delta_{p-1}}(x, y | \lambda)| \leq \text{second member of (11.7a);}$$

moreover, $\Omega^{\delta_1, \dots, \delta_{p-1}}(x, y | \lambda)$ will have the same descriptive properties as $\Omega^{\delta_1, \dots, \delta_p}(x, y | \lambda)$. Continuing in this manner along the lines of §4 we conclude that the results of Theorem 4.1 all hold for the kernel (which is generally of transfinite rank) $K^{\delta_1, \dots, \delta_{p-1}}(x, y)$. Continuing the reasoning of the type employed before, passing to the limit, it is established that results of such type hold for $K(x, y)$, itself. Finally, by transfinite induction the following Theorem is established.

THEOREM 11.1. *Let $K(x, y)$ be a kernel of class H_β , where β is any number of the first or second class (Definition 10.1). With respect to this kernel all the results (stated in appropriate form) of Theorem 4.1 will hold.*

Many more significant results for kernels H_β ($\beta > \omega$) can be obtained whenever with the kernel in question one may associate an operator L satisfying the following definition.

DEFINITION 11.1. *A linear operator $L_x(\xi | h(x))$ (ξ a parameter) will be said to be associated with $K(x, y) \in H_\beta$ ($\beta > \omega$) (cf. Definition 10.1) if*

$$(11.9) \quad L_x(\xi | K(x, y)) \in L_2 \quad (\text{in } y);$$

$$(11.10) \quad |L_x(\xi | K^{\delta_1, \dots, \delta_q}(x, y))| < \gamma(\xi | y),$$

with $\gamma(\xi | y) \in L_2$ (in y), the function $\gamma(\xi | y)$ being independent of $\delta_1, \dots, \delta_q$;

$$(11.11) \quad L_x(\xi | K^{\delta_1, \dots, \delta_q}(x, y)) \xrightarrow{\delta_q} L_x(\xi | K^{\delta_1, \dots, \delta_{q-1}}(x, y)) \\ \xrightarrow{\delta_{q-1}} L_x(\xi | K^{\delta_1, \dots, \delta_{q-2}}(x, y)) \rightarrow \dots \xrightarrow{\delta_1} L_x(\xi | K(x, y));$$

$$(11.12) \quad \lim_p L_x(\xi | f_p(x)) = L_x(\xi | f(x)) \quad [\text{when } f_p(x), \text{ included in } L_2, \rightarrow f(x) \text{ weakly}];$$

$$(11.13) \quad \int_a^b L_x(\xi | K^{\delta_1, \dots, \delta_q}(x, y)) \phi(y) dy = L_x \left(\xi \left| \int_a^b K^{\delta_1, \dots, \delta_q}(x, y) \phi(y) dy \right. \right)$$

for all $\phi(x) \in L_2$. Here $\delta_1, \dots, \delta_q$ are the numbers referred to in Definition 10.1.

Before proceeding further it is essential to give an example of a kernel of

transfinite rank, which is at the same time not of any finite rank and with which one can "associate" in the sense of Definition 11.1, an operator L . For this purpose consider the kernel $K(x, y)$ defined by (10.8)–(10.8c) (cf. text from (10.5) to (10.8c)); $K(x, y) \in H_{\omega+1}$ and $K(x, y)$ does not belong to any H_ν with $\nu < \omega$. The associated operator will be taken of the form

$$(11.14) \quad L_x(\xi | h(x)) = \int_0^1 G(\xi | x) h(x) dx,$$

where

$$(11.14a) \quad G(\xi | x) = 0 \quad (3/4 \leq x \leq 1),$$

$$(11.14b) \quad G(\xi | x) = c_i^{-1/2}(b_i^2 - x^2)^{1/2} w_i(\xi | x) \quad (\gamma_i = (a_i + b_i)/2 \leq x < b_i),$$

$$(11.14c) \quad G(\xi | x) = -G_i(\xi | a_i + b_i - x), \quad (a_i < x \leq \gamma_i),$$

the a_i and b_i being the numbers from (10.8b),

$$(11.14d) \quad 0 \leq w_i(\xi | x) \leq H, \quad w_i(\xi | x) \in L_1 \quad (\text{in } x; \gamma_i \leq x < b_i),$$

the function $w_i(\xi | x)$ being monotone non-increasing in x on (γ_i, b_i) .

By (11.14b) and (11.14d)

$$|G(\xi | x)|^2 \leq \frac{H^2}{c_i} (b_i^2 - \gamma_i^2) \quad (\gamma_i \leq x < b_i),$$

and, in view of the symmetry relation (11.14c),

$$(11.15) \quad |G(\xi | x)|^2 \leq \frac{H^2}{c_i} (b_i^2 - \gamma_i^2) \quad (a_i < x < b_i).$$

Thus

$$(11.15a) \quad \int_a^b |G(\xi | x)|^2 dx = \sum_i \int_{a_i}^{b_i} |G(\xi | x)|^2 dx \\ \leq H^2 \sum_i \frac{1}{c_i} (b_i^2 - \gamma_i^2)(b_i - a_i) = H^2 S.$$

Herewith we choose the c_i so that the series S of the last member in (11.15a) converges.* Accordingly,

$$(11.16) \quad G(\xi | x) \in L_2 \quad (\text{in } x; 0 \leq x \leq 1).$$

In consequence of (11.14), (10.8), and (10.8a),

$$(11.17) \quad L_x(\xi | K(x, y)) = \beta(\xi | y) + \alpha(\xi | y), \\ \beta(\xi | y) = g(y) \int_0^y G(\xi | x) dx, \quad \alpha(\xi | y) = \int_y^1 G(\xi | x) g(x) dx,$$

* This, obviously, it is always possible to do.

$g(x)$ being defined by (10.8c). By (10.8c)

$$(11.17a) \quad \beta(\xi | y) = 0 \quad (a_i < y < \gamma_i; 3/4 \leq y \leq 1).$$

On the other hand, for $\gamma_i \leq y < b_i$,

$$\int_0^y G(\xi | x) dx = \int_{a_i}^y G(\xi | x) dx + \sum^1 \int_{a_j}^{b_j} G(\xi | x) dx,$$

where the summation symbol is over the subscripts j , corresponding to the totality of all those numbers b_j for which $b_j < b_i$. Now γ_j is the mid-point of the interval (a_j, b_j) ; thus, on taking account of the symmetry relation (11.14c) (for j) we conclude that

$$(11.17b) \quad \int_a^{b_j} G(\xi | x) dx = 0.$$

Hence

$$\int_0^y G(\xi | x) dx = \int_{a_i}^y G(\xi | x) dx \quad (\gamma_i \leq y < b_i),$$

so that, by (11.14c), (11.14b), and in view of the monotone character of $w_i(\xi | x)$,

$$(11.17c) \quad \begin{aligned} \int_0^y G(\xi | x) dx &= \int_{a_i}^{2\gamma_i - y} G(\xi | u) du = - \int_y^{b_i} G(\xi | u) du \\ &= - c_i^{-1/2} \int_y^{b_i} (b_i^2 - u^2)^{1/2} w_i(\xi | u) du \quad (\gamma_i \leq y < b_i); \end{aligned}$$

$$(11.17d) \quad \left| \int_0^y G(\xi | x) dx \right| \leq c_i^{-1/2} (b_i^2 - y^2)^{1/2} w_i(\xi | y) (b_i - y) \quad (\gamma_i \leq y < b_i).$$

Thus, in consequence of (11.17), (10.8c), and (11.17d), it is inferred that

$$|\beta(\xi | y)| \leq w_i(\xi | y) (b_i - \gamma_i) \quad (\gamma_i \leq y < b_i),$$

whence by (11.14d)

$$|\beta(\xi | y)| \leq h(b_i - \gamma_i) < h \quad (\gamma_i \leq y < b_i),$$

which, together with (11.17a), implies that

$$(11.18) \quad |\beta(\xi | y)| < h \quad (0 \leq y \leq 1).$$

For the function $\alpha(\xi | y)$, involved in (11.17), one has

$$|\alpha(\xi | y)| \leq \alpha(\xi) = \int_0^1 |G(\xi | x)| g(x) dx = \sum_i \int_{a_i}^{b_i} |G(\xi | x)| g(x) dx,$$

and, in consequence of (10.8c), (11.14b), (11.14d),

$$\begin{aligned} |\alpha(\xi | y)| &\leq \alpha(\xi) = \sum_i \int_{\gamma_i}^{b_i} |G(\xi | x)| \frac{c_i^{1/2}}{(b_i^2 - x^2)^{1/2}} dx \\ (11.18a) \qquad &= \sum_i \int_{\gamma_i}^{b_i} w_i(\xi | x) dx < H \end{aligned}$$

for $0 \leq y \leq 1$. Consideration of (11.17), (11.18), (11.18a) leads to the conclusion that *condition (11.9) of Definition 11.1 holds for the case under consideration.*

Consider the related kernel $K^{\delta_1, \dots, \delta_{j+2}}(x, y)$ (cf. (10.9)–(10.9b)). One has

$$(11.19) \quad L_x(\xi | \gamma^{\delta_1, \dots, \delta_{j+2}}(x, y)) = \beta^{\delta_1, \dots, \delta_{j+2}}(\xi | y) + \alpha^{\delta_1, \dots, \delta_{j+2}}(\xi | y),$$

$$\beta^{\delta_1, \dots, \delta_{j+2}}(\xi | y) = g^{\delta_1, \dots, \delta_{j+2}}(y) \int_0^y G(\xi | x) dx, \quad (11.19a)$$

$$\alpha^{\delta_1, \dots, \delta_{j+2}}(\xi | y) = \int_y^1 G(\xi | x) g^{\delta_1, \dots, \delta_{j+2}}(x) dx.$$

We have, as can be seen from (10.9b),

$$0 \leq |g^{\delta_1, \dots, \delta_{j+2}}(x)| \leq g(x), \quad \lim_{\delta_1} \lim_{\delta_2} \dots \lim_{\delta_{j+2}} g^{\delta_1, \dots, \delta_{j+2}}(x) = g(x).$$

Thus, by (11.17) and (11.18),

$$\begin{aligned} (11.19b) \quad |\beta^{\delta_1, \dots, \delta_{j+2}}(\xi | y)| &\leq g(x) \left| \int_0^y G(\xi | x) dx \right| = |\beta(\xi | y)| < H \\ &\qquad\qquad\qquad (0 \leq y \leq 1). \end{aligned}$$

Also, in view of (11.18a),

$$\begin{aligned} (11.19c) \quad |\alpha^{\delta_1, \dots, \delta_{j+2}}(\xi | y)| &\leq \int_0^1 |G(\xi | x)| g^{\delta_1, \dots, \delta_{j+2}}(\xi | x) dx \\ &\leq \int_0^1 |G(\xi | x)| g(x) dx = \alpha(\xi) < H. \end{aligned}$$

By virtue of (11.19), (11.19b), (11.19c) it can be asserted that *condition (11.10) of Definition 11.1 holds, with $\gamma(\xi | y) = 2H$.**

To demonstrate the first relation (11.11) it will suffice to establish

$$\begin{aligned} (11.20) \quad \lim_{\delta_{j+2}} \beta^{\delta_1, \dots, \delta_{j+2}}(\xi | y) &= \beta^{\delta_1, \dots, \delta_{j+1}}(\xi | y), \\ \lim_{\delta_{j+2}} \alpha^{\delta_1, \dots, \delta_{j+2}}(\xi | y) &= \alpha^{\delta_1, \dots, \delta_{j+1}}(\xi | y), \end{aligned}$$

* Or, more precisely, with $\gamma(\xi | y) = |\beta(\xi | y)| + \alpha(\xi)$.

$$\begin{aligned}
 (11.20a) \quad \beta^{\delta_1, \dots, \delta_{j+1}}(\xi | y) &= g^{\delta_1, \dots, \delta_{j+1}}(y) \int_0^y G(\xi | x) dx, \\
 \alpha^{\delta_1, \dots, \delta_{j+1}}(\xi | y) &= \int_y^1 G(\xi | x) g^{\delta_1, \dots, \delta_{j+1}}(x) dx
 \end{aligned}$$

where

$$\begin{aligned}
 (11.20b) \quad g^{\delta_1, \dots, \delta_{j+1}}(x) &= 0 & [x \text{ in } \Delta^1(\delta_1) + \dots + \Delta^{j+1}(\delta_{j+1})], \\
 g^{\delta_1, \dots, \delta_{j+1}}(x) &= g(x) & [x \text{ in } (0, 1) - (\Delta^1(\delta_1) + \dots + \Delta^{j+1}(\delta_{j+1}))].
 \end{aligned}$$

Now, the first relation (11.20) follows from (11.19a), since

$$(11.21) \quad \lim_{\delta_{j+2}} g^{\delta_1, \dots, \delta_{j+2}}(y) = g^{\delta_1, \dots, \delta_{j+1}}(y).$$

The remaining part of (11.20) will hold if

$$(11.22) \quad \lim_{\delta_{j+2}} \int_y^1 G(\xi | x) g^{\delta_1, \dots, \delta_{j+2}}(x) dx = \int_y^1 G(\xi | x) g^{\delta_1, \dots, \delta_{j+1}}(x) dx.$$

In view of the inequality subsequent to (11.9a)

$$|G(\xi | x) g^{\delta_1, \dots, \delta_{j+2}}(x)| \leq |G(\xi | x)| g(x).$$

The last number, here, is contained in L_1 (in x), as can be inferred from the existence of the function $\alpha(\xi)$, introduced subsequent to (11.18). On the other hand, in consequence of (11.21) the limit (as $\delta_{j+2} \rightarrow 0$) of the integrand in the first member of (11.22) converges to the integrand of the second member. Thus the passage to the limit under the integral sign, indicated in (11.22), is justifiable. The first condition (11.11) accordingly holds. *All the other conditions (11.11) can be demonstrated in succession following the indicated procedure.*

In view of (11.14) justification of (11.12) amounts to that of

$$(11.23) \quad \lim_{\delta_{j+2}} \int_0^1 G(\xi | x) f_{\delta_{j+2}}(x) dx = \int_0^1 G(\xi | x) f(x) dx \quad (f_{\delta_{j+2}}(x) \rightarrow f(x) \text{ weakly}).$$

This relationship, however, holds in consequence of (11.16) and of Theorem 1.4.

Finally, demonstration of the condition (11.13) for the case under consideration is effected by noting that, in view of (10.9d), the change of order of integration, involved in the relationship

$$\begin{aligned}
 (11.24) \quad & \int_{y=0}^1 \left[\int_{x=0}^1 G(\xi | x) K^{\delta_1, \dots, \delta_{j+2}}(x, y) dx \right] \phi(y) dy \\
 &= \int_{x=0}^1 G(\xi | x) \left[\int_{y=0}^1 K^{\delta_1, \dots, \delta_{j+2}}(x, y) \phi(y) dy \right] dx \quad (\phi(y) \in L_2),
 \end{aligned}$$

is justifiable.

Thus, $K(x, y)$, as defined by (10.8)–(10.8c), belongs to $H_{\omega+1}$ and does not belong to any H_ν with $\nu < \omega$; moreover, “associated” (in the sense of Definition 11.1) with $K(x, y)$ there is an operator L of the form (11.14)–(11.14c).

Let $K(x, y)$ be any kernel of a transfinite class H_β (any β of 2d class) such that with $K(x, y)$ there is “associated” an operator L . Then it is of interest to study the equation

$$(11.25) \quad L_x(\xi | \phi(x)) - \lambda \int_a^b L_x(\xi | K(x, y))\phi(y)dy = L_x(\xi | f(x)).$$

For this purpose it is advantageous to consider the auxiliary equation

$$(11.25a) \quad L_x(\xi | \phi^{\delta_1, \dots, \delta_q}(x)) - \lambda \int_a^b L_x(\xi | K^{\delta_1, \dots, \delta_q}(x, y))\phi^{\delta_1, \dots, \delta_q}(y)dy = L_x(\xi | f(x));$$

here the δ_ν ($\nu=1, \dots, q$) are the numbers involved in Definition 10.1. Of importance is also the homogeneous equation

$$(11.26) \quad L_x(\xi | \phi(x)) - \lambda \int_a^b L_x(\xi | K(x, y))\phi(y)dy = 0.$$

Using the results of Theorems 4.1, 5.1, 7.1, 11.1, partly by direct methods and partly by transfinite induction and following the lines which were employed previously, we arrive at the following theorem.

THEOREM 11.2. For kernels $K(x, y) \in H_\beta$ (any β of 2d class) [classes H_β are specified in Definition 10.1], for which there exist “associated” operators L (Definition 11.1), all the results of Theorems 5.1, 7.1 will hold, if appropriately stated with respect to the equations (11.25), (11.26).

In conclusion we shall point out that if $K(x, y) \in H_\beta$ (β possibly transfinite; Definition 10.1) and if

$$(11.27) \quad \int_a^b (K(x, y) - K(x^1, y))^2 dy = g(x, x^1)$$

exists for x and x^1 in $(a, b) - E$ (E the set of §9), $g(x, x^1)$ being continuous in x and x^1 in $(a, b) - E$, then the following will hold:

$$(11.27a) \quad \lim_{\delta_{q,r}} \frac{\partial}{\partial x} \Omega^{\delta_1, \dots, \delta_{q,r}}(x, y | \lambda) = \frac{\partial}{\partial x} \Omega^{\delta_1, \dots, \delta_{q-1}}(x, y | \lambda), \dots,$$

$$\lim_{\delta_{1,r}} \frac{\partial}{\partial x} \Omega^{\delta_1, r}(x, y | \lambda) = \frac{\partial}{\partial x} \Omega(x, y | \lambda);$$

$$(11.27b) \quad \lim_{\delta_{q,r}} \frac{\partial}{\partial y} \Omega^{\delta_1, \dots, \delta_q, r}(x, y | \lambda) = \frac{\partial}{\partial y} \Omega^{\delta_1, \dots, \delta_{q-1}}(x, y | \lambda), \dots;$$

$$(11.27c) \quad \lim_{\delta_{q,r}} \frac{\partial^2}{\partial x \partial y} \Omega^{\delta_1, \dots, \delta_q, r}(x, y | \lambda) = \frac{\partial^2}{\partial x \partial y} \Omega^{\delta_1, \dots, \delta_{q-1}}(x, y | \lambda), \dots,$$

$$\lim_{\delta_{1,r}} \frac{\partial^2}{\partial x \partial y} \Omega^{\delta_1, r}(x, y | \lambda) = \frac{\partial^2}{\partial x \partial y} \Omega(x, y | \lambda) \quad [= \theta(x, y | \lambda)],$$

provided the $\delta_{v,r}$ are suitably chosen. Convergence in the above relations will be uniform in any closed subset of $0 \leq x, y \leq b$, which has no points in common with the lines $x = I_\nu, y = I_\nu$ ($\nu = 1, 2, \dots$).*

To prove this fact we need only to replace in (C, pp. 145, 146) θ_δ and K_δ by $\theta^{\delta_1, \dots, \delta_q}$ and $K^{\delta_1, \dots, \delta_q}$, respectively. This will yield

$$|\theta^{\delta_1, \dots, \delta_q}(x, y | \lambda)|$$

$$< \left[\frac{1}{(b^1 - a^1)^{1/2}} + \frac{|\lambda|}{b^1 - a^1} \int_{a^1}^{b^1} \left[\int_a^b [K(x, s) - K(x^1, s)]^2 ds \right]^{1/2} dx^1 \right]$$

$$\cdot \left[\frac{1}{(b^1 - a^1)^{1/2}} + \frac{|\lambda|}{b^1 - a^1} \int_{a^1}^{b^1} \left[\int_a^b [K(y, s) - K(x^1, s)]^2 ds \right]^{1/2} dx^1 \right]$$

$$[(a^1, b^1) \text{ a closed subinterval of } (a, b) - E],$$

$$|\theta^{\delta_1, \dots, \delta_q}(x, y | \lambda) - \theta^{\delta_1, \dots, \delta_q}(x^1, y | \lambda)|$$

$$\leq |\lambda| \left[\int_a^b |K(x, s) - K(x^1, s)|^2 ds \right]^{1/2}$$

$$\cdot \left[\frac{1}{(b^1 - a^1)^{1/2}} + \frac{|\lambda|}{b^1 - a^1} \int_{a^1}^{b^1} \left[\int_a^b [K(y, s) - K(t, s)]^2 ds \right]^{1/2} dt \right].$$

Using these inequalities, the stated result will follow with the aid of consecutive applications of the "Compactness Theorem" (§1).

12. Non-symmetric kernels. Let $K(x, y)$ be a kernel not necessarily symmetric. We let

$$(12.1) \quad E_1, E_2$$

denote reducible sets on (a, b) , each of the description given in the beginning of §9, with

$$(12.1a) \quad E^{\beta_1} = 0, \quad E^{\beta_2} = 0,$$

where β_1, β_2 are non-limit numbers of the 1st or 2d class and the sets

$$(12.1b) \quad E^{\beta_1-1}, \quad E^{\beta_2-1}$$

each have some points. In accordance with Theorem 9.1 the set E_1 will be

* E consists of the points represented by the numbers I_ν .

"covered" by sets of intervals

$$(12.2) \quad \Delta^{1,v}(\delta_{1,v}) \quad (v = 1, 2, \dots, q_1 < \omega)$$

and the set E_2 will be "covered" by sets of intervals

$$(12.2a) \quad \Delta^{2,v}(\delta_{2,v}) \quad (v = 1, 2, \dots, q_2 < \omega).$$

DEFINITION 12.1. A non-symmetric kernel $K(x, y)$ will be said to belong to the class $H(\beta_1, \beta_2)$ if, with the text from (12.1) to (12.2a) in view, the following is true for all "admissible" positive values $\delta_{1,v}, \delta_{2,v}$:

$$(12.3) \quad G(x, y) = K_{\delta_{1,1}, \delta_{1,2}, \dots, \delta_{1,q_1}, \delta_{2,1}, \delta_{2,2}, \dots, \delta_{2,q_2}}(x, y) \in L_2 \quad (\text{in } x, y; \text{ for } a \leq x, y \leq b).$$

Here

$$(12.3a) \quad G(x, y) = 0 \quad [x \text{ in } \sum_1^{q_1} \Delta^{1,v}(\delta_{1,v}), a \leq y \leq b];$$

$$(12.3b) \quad G(x, y) = 0 \quad [y \text{ in } \sum_1^{q_2} \Delta^{2,v}(\delta_{2,v}), a \leq x \leq b];$$

$$(12.3c) \quad G(x, y) = K(x, y) \quad [\text{at all other points of } a \leq x, y \leq b].$$

Using the well known method of Schmidt one may associate with a non-symmetric kernel a pair of integral equations, whose kernels are symmetric. However, we shall find it more convenient to employ the device of Pérès* and, thus, associate with our kernel $T(x, y)$ a single symmetric kernel $T(x, y)$ defined as follows

$$(12.4) \quad \begin{aligned} T(x, y) &= 0 & (a < x, y < b), \\ T(x, y) &= 0 & (b < x, y < 2b - a), \\ T(x, y) &= K(x, y + a - b) & (a < x < b, b < y < 2b - a), \\ T(x, y) &= K(y, x + a - b) & (b < x < 2b - a, a < y < b). \end{aligned}$$

Inasmuch as $K(x, y) \in H(\beta_1, \beta_2)$ (Definition 12.1), one clearly has (Definition 10.1)

$$(12.5) \quad T(x, y) \in H_\beta,$$

where β is the greater one of the numbers β_1, β_2 . The set E , which according to Definition 10.1 is used in the description of a kernel of class H_β , consists (in the case on hand) of the points of E_1 on (a, b) and of the points $b + I_{2,v}$ [$v = 1, 2, \dots$; the $I_{2,v}$ represent points of E_2].

We apply to $T(x, y)$ the results of the previous sections; this will lead to conclusions with respect to the given non-symmetric kernel $K(x, y)$.

* Volterra and Pérès, loc. cit., p. 306.

LIMITS OF A DISTRIBUTION FUNCTION DETERMINED BY ABSOLUTE MOMENTS AND INEQUALITIES SATISFIED BY ABSOLUTE MOMENTS*

BY
ABRAHAM WALD†

1. Introduction. Denote by X a chance variable and by $P(\alpha < X < \beta)$ the probability that $\alpha < X < \beta$. Similarly denote by $P(X < \beta)$ the probability that $X < \beta$ and by $P(X = \beta)$ the probability that $X = \beta$. For any positive integer r the expected value $E|X - x_0|^r$ of $|X - x_0|^r$ is called the absolute moment of order r about x_0 , where x_0 denotes a certain real value. If the absolute moments $M_{i_1} = E|X - x_0|^{i_1}, \dots, M_{i_j} = E|X - x_0|^{i_j}$ of a chance variable X are given (and no further data about X are known), then we shall say for any positive number d that a_d is the sharp lower limit of $P(-d < X - x_0 < d)$ if the following two conditions are fulfilled:

(1) For each chance variable Y for which $E|Y - x_0|^{i_\nu} = E|X - x_0|^{i_\nu}$ ($\nu = 1, \dots, j$) the inequality $P(-d < Y - x_0 < d) \geq a_d$ holds.

(2) To each $\epsilon > 0$ a chance variable Y can be given such that $E|Y - x_0|^{i_\nu} = E|X - x_0|^{i_\nu}$ ($\nu = 1, \dots, j$) and $P(-d < Y - x_0 < d) < a_d + \epsilon$.

In other words, a_d is the greatest lower bound of the probabilities $P(-d < Y - x_0 < d)$ formed for all chance variables Y for which the i_ν th absolute moment about x_0 is equal to the i_ν th absolute moment of X about x_0 ($\nu = 1, \dots, j$).

Similarly we shall say that b_d is the sharp upper limit of $P(-d < X - x_0 < d)$ if b_d is the least upper bound of the probabilities $P(-d < Y - x_0 < d)$ formed for all chance variables Y for which the i_ν th absolute moment about x_0 is equal to the i_ν th absolute moment of X about x_0 ($\nu = 1, \dots, j$).

In this paper we shall give the solution of the following two problems:

PROBLEM 1. *The absolute moments of the order i_1, \dots, i_j of a chance variable X are given about the point x_0 , where i_1, \dots, i_j denote any positive integers. It is required to determine the sharp lower and sharp upper limit of $P(-d < X - x_0 < d)$ for any positive value d .*

PROBLEM 2. *A real value x_0 and a system of j positive integers i_1, \dots, i_j are given. What are the necessary and sufficient conditions which must be satisfied*

* Presented to the Society, February 25, 1939; received by the editors February 15, 1939.

† This research was done under a grant-in-aid from the Carnegie Corporation of New York.

by j positive numbers a_1, \dots, a_j that a chance variable X exists for which the i th moment about x_0 is equal to a_i ($i=1, \dots, j$)?

The solution of Problem 1 is a generalization of the inequality of Markoff. In fact, the inequality of Markoff can be written as follows:

$$(1) \quad P(-d < X - x_0 < d) \geq 1 - M_r/d^r,$$

where d denotes an arbitrary positive value and M_r denotes the r th absolute moment of X about x_0 . As is well known, the inequality (1) cannot be improved for $d \geq M_r^{1/r}$, that is to say that $1 - M_r/d^r$ is the sharp lower limit of $P(-d < X - x_0 < d)$ for $d \geq M_r^{1/r}$. The generalization in our Problem 1 consists in the circumstance that instead of a single moment M_r , we consider a finite number of moments M_{i_1}, \dots, M_{i_j} , and besides the sharp lower limit of $P(-d < X - x_0 < d)$ also its sharp upper limit is to be determined. The inequality (1) is called for $r=2$ also the inequality of Tshebysheff.

Some results concerning the case when two moments M_r and M_s are given, have been obtained by different authors. A. Guldberg* gave the following formula:

$$(2) \quad P(|X - x_0| < \lambda M_r^{1/r}) \geq 1 - \frac{1}{\lambda^s} \left(\frac{M_s^{1/s}}{M_r^{1/r}} \right)^s.$$

If we substitute $2k$ for s , and 2 for r , we get the inequality of K. Pearson.† By other substitutions we get the formula of E. Lurquin.‡ It is easy to show that the limit given in (2) is not sharp.

P. Cantelli§ gave a formula in case that $s=2r$. His formula can be written as follows:

$$(3a) \quad \text{If } M_r/d^r \leq M_{2r}/d^{2r}, \text{ then } P(|X - x_0| < d) \geq 1 - M_r/d^r.$$

$$(3b) \quad \text{If } M_r/d^r > M_{2r}/d^{2r}, \text{ then}$$

$$P(|X - x_0| < d) \geq 1 - \frac{M_{2r} - M_r^2}{(d^r - M_r)^2 + M_{2r} - M_r^2}.$$

The writer of this article gave in a previous paper|| some results concerning the general case and the sharp lower limit of $P(-d < X - x_0 < d)$ if two

* A. Guldberg, Comptes Rendus de l'Académie des Sciences, Paris, vol. 175, p. 679.

† K. Pearson, Biometrika, vol. 12 (1918-1919).

‡ E. Lurquin, Comptes Rendus de l'Académie des Sciences, Paris, vol. 175, p. 681.

§ Cantelli's formula and its demonstration are given in the book of M. Fréchet, *Recherches Théoriques Modernes sur la Théorie des Probabilités*, Paris, 1937, pp. 123-126.

|| A. Wald, *A generalization of Markoff's inequality*, Annals of Mathematical Statistics, December, 1938.

moments M_r and M_s are given, where r and s denote arbitrary positive integers. If $s = 2r$ the formula reduces to Cantelli's formula.

In case of consecutive algebraic moments, that is to say, if M_1, \dots, M_j are given and $M_i = E(X - x_0)^i$ ($i = 1, \dots, j$), Tshebysheff determined the sharp lower and sharp upper limit of the distribution function $P(X < d)$. These inequalities are called Tshebysheff's inequalities. The first proof of these inequalities was given by Markoff in 1884 and the same proof was discovered almost at the same time by Stieltjes.*

The solution of Problem 2 is well known† if i_1, \dots, i_j are consecutive integers, that is to say, if $i_\nu = \nu$ ($\nu = 1, \dots, j$) and if a_ν ($\nu = 1, \dots, j$) is the ν th algebraic moment, that is to say, $a_\nu = E(X - x_0)^\nu$. In this paper we shall give the solution for absolute moments and for arbitrary positive integers i_1, \dots, i_j .

2. Reduction of the problem to the case of nonnegative chance variables.

We shall call a chance variable X nonnegative if $P(X < 0) = 0$. Since the moments of the nonnegative chance variable $Y = |X - x_0|$ about the origin are equal to the absolute moments of X about x_0 and since

$$P(Y < d) = P(-d < X - x_0 < d),$$

the following proposition holds true:

PROPOSITION 1. Denote by M_{i_1}, \dots, M_{i_j} the absolute moments of order i_1, \dots, i_j of a certain chance variable X about the point x_0 . There exists a nonnegative chance variable Y such that the i_ν th moment of Y about the origin is equal to M_{i_ν} ($\nu = 1, \dots, j$). The greatest lower (least upper) bound of the probabilities $P(-d < Z - x_0 < d)$ is equal to the greatest lower (least upper) bound of the probabilities $P(Z' < d)$, where $P(-d < Z - x_0 < d)$ is formed for all chance variables Z for which the i_ν th absolute moment about x_0 is equal to M_{i_ν} and $P(Z' < d)$ is formed for all nonnegative chance variables Z' for which the i_ν th moment about the origin is equal to M_{i_ν} ($\nu = 1, \dots, j$).

On account of Proposition 1 we can restrict ourselves to the consideration of nonnegative chance variables and of the moments about the origin. Throughout the following developments we shall understand by a chance variable a nonnegative chance variable and by moments the moments about the origin.

3. Some definitions and propositions. Let us begin with some definitions.

* See, for instance, J. Uspensky, *Introduction to Mathematical Probability*, New York, McGraw-Hill, 1937, pp. 373-380.

† See, for instance, R. von Mises, *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*, Deuticke, Leipzig, 1931, pp. 247-248.

DEFINITION 1. A chance variable X is said to be an arithmetic chance variable, if there exist a finite system of different numbers x_1, \dots, x_k such that $\sum_{i=1}^k P(X=x_i) = 1$.

DEFINITION 2. A chance variable X for which k different positive values x_1, \dots, x_k exist such that $P(X=x_i) > 0$ ($i=1, \dots, k$) and $\sum_{i=1}^k P(X=x_i) = 1$, is called an arithmetic chance variable of the degree k .

DEFINITION 3. A chance variable X is said to be an arithmetic chance variable of the degree $k+1/2$ if $P(X=0) > 0$ and if there exist k different positive values x_1, \dots, x_k such that $P(X=x_i) > 0$ ($i=1, \dots, k$) and $\sum_{i=1}^k P(X=x_i) + P(X=0) = 1$.

DEFINITION 4. Denote by M_{i_1}, \dots, M_{i_j} the moments of the order i_1, \dots, i_j of a certain chance variable X . A chance variable Y is said to be characteristic relative to M_{i_1}, \dots, M_{i_j} if the i th moment of Y is equal to M_{i_v} ($v=1, \dots, j$) and Y is an arithmetic chance variable of the degree less than or equal to $(j+1)/2$. A characteristic chance variable is said to be degenerate if its degree is less than $(j+1)/2$.

DEFINITION 5. We shall say that the numbers M_{i_1}, \dots, M_{i_j} can be realized as moments of the order i_1, \dots, i_j if there exists a chance variable X such that the i th moment of X is equal to M_{i_v} ($v=1, \dots, j$).

DEFINITION 6. A function $f(x)$ defined for all real values x is said to change its sign at the point $x=\alpha$ if the following conditions are fulfilled:

- (1) If $f(x) = 0$ for all values $x < \alpha$, then any open interval containing α must contain at least one value α' such that $f(\alpha)f(\alpha') < 0$.
- (2) If $f(x)$ is not identically zero for $x < \alpha$, then any open interval which contains α and a point $\beta < \alpha$ for which $f(\beta) \neq 0$, must also contain two points α_1 and α_2 such that $\alpha_1 \leq \alpha$, $\alpha_2 \geq \alpha$ and $f(\alpha_1)f(\alpha_2) < 0$.

By the number of changes in sign of $f(x)$ we shall understand the number of points at which $f(x)$ changes its sign. Similarly we shall understand by the number of changes in sign in an (open or closed) interval A , the number of points of A at which $f(x)$ changes its sign.

It is easy to prove that if $f(\alpha_1)f(\alpha_2) < 0$ then there exists at least one point of the closed interval $[\alpha_1, \alpha_2]$ at which $f(x)$ changes its sign. In order to prove this, let us assume that $\alpha_1 < \alpha_2$ and denote by α the greatest lower bound of all values γ of the interval $[\alpha_1, \alpha_2]$ for which $f(\alpha_1)f(\gamma) < 0$. It is obvious that $\alpha_1 \leq \alpha \leq \alpha_2$. We shall show that $f(x)$ changes its sign at α . If $\alpha = \alpha_1$ then from the definition of α it follows that any open interval containing α contains also a point α' such that

$$f(\alpha_1)f(\alpha') = f(\alpha)f(\alpha') < 0.$$

Hence $f(x)$ changes its sign at α . If $\alpha > \alpha_1$ then for any value $\delta \geq \alpha_1$ and less than α , $f(\delta)$ has the same sign as $f(\alpha_1)$ or is equal to zero. From this fact it follows easily that any open interval which contains α and a value $\beta < \alpha$ for which $f(\beta) \neq 0$, contains also two points β_1 and β_2 such that $\beta_1 \leq \alpha$, $\beta_2 \geq \alpha$ and $f(\beta_1)f(\beta_2) < 0$. Hence $f(x)$ changes its sign at α in any case.

If $f(x)$ does not change its sign at any point of the (open or closed) interval I , then $f(\alpha)f(\beta) \geq 0$ for any two points α, β of I . In fact if I should contain two points α, β such that $f(\alpha)f(\beta) < 0$, then $[\alpha, \beta]$ and therefore also I must contain a point γ at which $f(x)$ changes its sign, in contradiction to our assumption.

We shall prove now the

PROPOSITION 2. *If X denotes an arithmetic chance variable of degree k and Y denotes an arbitrary chance variable, then the number of changes in sign of $D(x) = P(X < x) - P(Y < x)$ is less than or equal to $2k - 1$.*

Let us first consider the case that k is integral. In this case there are k different positive values $\alpha_1, \dots, \alpha_k$ such that $P(X = \alpha_i) > 0$ ($i = 1, \dots, k$) and $\sum_{i=1}^k P(X = \alpha_i) = 1$. It is obvious that at most one change in sign of $D(x)$ can take place in the interior of the interval $I_i = [\alpha_i, \alpha_{i+1}]$ ($i = 1, \dots, k-1$). Besides the changes in sign in the interior of the intervals I_1, \dots, I_{k-1} a change in sign can only occur at the points $\alpha_1, \dots, \alpha_k$. Hence the total number of changes in sign cannot exceed $(k-1) + k = 2k - 1$.

If $k = k' + 1/2$, where k' denotes a nonnegative integer, then $P(X = 0) > 0$ and there exist k' different positive values $\alpha_1, \dots, \alpha_{k'}$ such that $P(X = \alpha_i) > 0$ ($i = 1, \dots, k'$) and $\sum_{i=1}^{k'} P(X = \alpha_i) + P(X = 0) = 1$. Let us denote the point 0 by α_0 . It is obvious that in the interior of the interval $I_i = [\alpha_i, \alpha_{i+1}]$ ($i = 0, 1, \dots, k'-1$) at most one change in sign of $D(x)$ can take place. Further changes in sign can occur only at the points $\alpha_1, \dots, \alpha_{k'}$. Hence the total number of changes in sign cannot exceed $2k' = 2k - 1$.

PROPOSITION 3. *If X and Y denote two arithmetic chance variables of degree less than or equal to k , then the number of changes in sign of $D(x) = P(X < x) - P(Y < x)$ is less than or equal to $2k - 2$.*

First let us consider the case that both chance variables X and Y are of the degree k . If k is a positive integer, then there exist two systems of k positive values $\alpha_1, \dots, \alpha_k$ and β_1, \dots, β_k such that

$$\sum_{i=1}^k P(X = \alpha_i) = \sum_{i=1}^k P(Y = \beta_i) = 1.$$

We may assume that $\alpha_1 \leq \beta_1$. (If it happens that $\beta_1 < \alpha_1$, we can change the notation.) Hence $D(x)$ has no change in sign at the point α_1 . Since in the in-

terior of the interval $[\alpha_i, \alpha_{i+1}]$ ($i=1, \dots, k-1$) at most one change in sign can take place and further changes in sign can occur only at the points $\alpha_2, \alpha_3, \dots, \alpha_k$, the total number of changes in sign cannot exceed $(k-1) + (k-1) = 2k-2$. If $k=k'+1/2$, where k' denotes a nonnegative integer, then $P(X=0) > 0$, $P(Y=0) > 0$ and there exist two systems of k' positive numbers $\alpha_1, \dots, \alpha_{k'}; \beta_1, \dots, \beta_{k'}$ such that $P(X=\alpha_i) > 0$, $P(Y=\beta_i) > 0$ ($i=1, \dots, k'$) and

$$\sum_{i=1}^{k'} P(X=\alpha_i) + P(X=0) = \sum_{i=1}^{k'} P(Y=\beta_i) + P(Y=0) = 1.$$

We may assume that $\alpha_1 \leq \beta_1$. It is obvious that $D(x)$ has no change in sign in the interior of the interval $[0, \alpha_1]$. Since $D(x)$ has at most one change in sign in the interior of the interval $[\alpha_i, \alpha_{i+1}]$ ($i=1, \dots, k'-1$), and since further changes in sign can occur only at the points $\alpha_1, \dots, \alpha_{k'}$, the total number of changes in sign cannot exceed $2k'-1 = 2k-2$. Hence Proposition 3 is proved if X and Y are of the degree k .

Let us now consider the case that X or Y or both are of degree less than k . Let for instance the degree of X be less than k . Hence the degree of X is less than or equal to $k-1/2$ and therefore on account of Proposition 2 the number of changes in sign of $D(x)$ cannot exceed $2(k-1/2)-1 = 2k-2$.

PROPOSITION 4. *If X and Y denote two arithmetic chance variables of degree less than or equal to $k > 1$ and if there exists a positive number α such that $P(X=\alpha) > 0$ and $P(Y=\alpha) > 0$, then the number of changes in sign of $D(x) = P(X < x) - P(Y < x)$ is less than or equal to $2k-3$.*

We may assume that $P(X < \alpha) \leq P(Y < \alpha)$. Consider first the case that $P(Y < \alpha) > 0$ and denote by α' the greatest value less than α for which $P(Y=\alpha') > 0$. It is obvious that $D(x)$ has no change in sign in the interior of the interval $[\alpha', \alpha]$. If $D(x)$ is identically zero in the interior of $[\alpha', \alpha]$, then $D(x)$ has no change in sign at α' . If $D(x)$ is not identically zero in the interior of $[\alpha', \alpha]$ and if $P(X \leq \alpha) \leq P(Y \leq \alpha)$, then $D(x)$ has no change in sign at α . Finally if $P(X \leq \alpha) > P(Y \leq \alpha)$ and α'' denotes the smallest value greater than α for which $P(Y=\alpha'') > 0$, then $D(x)$ has no change in sign in the interior of the interval $[\alpha, \alpha'']$. Hence in any case the number of changes in sign of $D(x)$ cannot exceed $(2k-1)-2 = 2k-3$. Now we have to prove Proposition 4 if $P(Y < \alpha) = 0$. Since $P(X < \alpha) \leq P(Y < \alpha) = 0$, $D(x)$ has no change in sign at α . If $P(X \leq \alpha) = P(Y \leq \alpha) = 1$, then $D(x)$ has no change in sign at all and Proposition 4 is proved. We have to consider only the case that at least one of the values $P(X \leq \alpha)$, $P(Y \leq \alpha)$ is less than 1. Let us assume that $P(X \leq \alpha) \geq P(Y \leq \alpha)$. The probability $P(Y \leq \alpha)$ must be less than 1,

since otherwise also $P(X \leq \alpha)$ would be equal to 1, in contradiction to our assumption. Denote by β the smallest value greater than α for which $P(Y = \beta) > 0$. Then $D(x)$ has obviously no change in sign in the interior of $[\alpha, \beta]$ and therefore the total number of changes in sign cannot exceed $2k - 3$. If $P(X \leq \alpha) < P(Y \leq \alpha)$, then denote by β the smallest value greater than α for which $P(X = \beta) > 0$. The function $D(x)$ has no change in sign in the interior of $[\alpha, \beta]$ and therefore also in this case the total number of changes in sign of $D(x)$ cannot exceed $2k - 3$.

PROPOSITION 5. *Denote by X and Y two chance variables. If the i ,th moment ($\nu = 1, \dots, j$) of X is finite and equal to the i ,th moment of Y , then $D(x) = P(X < x) - P(Y < x)$ must have at least j changes in sign, unless $D(x)$ is identically zero.*

Denote $P(X < x)$ by $V_1(x)$ and $P(Y < x)$ by $V_2(x)$. Since the i ,th moment of X is equal to the i ,th moment of Y ($\nu = 1, \dots, j$), the Stieltjes integral

$$(4) \quad I = \int_0^\infty (a_1 x^{i_1} + \dots + a_j x^{i_j}) d[V_1(x) - V_2(x)] = 0$$

for arbitrary real values a_1, \dots, a_j . Denote the integral

$$\int_0^\lambda (a_1 x^{i_1} + \dots + a_j x^{i_j}) d[V_1(x) - V_2(x)]$$

by I_λ . It is obvious that

$$(5) \quad \lim_{\lambda \rightarrow \infty} I_\lambda = I = 0.$$

We get by integration by parts

$$(6) \quad \begin{aligned} I_\lambda &= (a_1 \lambda^{i_1} + \dots + a_j \lambda^{i_j}) [V_1(\lambda) - V_2(\lambda)] \\ &\quad - \int_0^\lambda (i_1 a_1 x^{i_1-1} + \dots + i_j a_j x^{i_j-1}) [V_1(x) - V_2(x)] dx. \end{aligned}$$

Now we shall show that

$$(7) \quad \lim_{\lambda \rightarrow \infty} \lambda^{i_\nu} [V_1(\lambda) - V_2(\lambda)] = 0, \quad \nu = 1, \dots, j.$$

Since

$$\lambda^{i_r} [V_1(\lambda) - V_2(\lambda)] = \lambda^{i_r} [1 - V_2(\lambda)] - \lambda^{i_r} [1 - V_1(\lambda)]$$

we have only to show that

$$(8) \quad \lim_{\lambda \rightarrow \infty} \lambda^{i_r} [1 - V_r(\lambda)] = 0, \quad r = 1, 2.$$

It is obvious that for any $\lambda > 0$

From Propositions 2 and 5 follows easily

PROPOSITION 6. *If X denotes an arithmetic chance variable of degree k and Y denotes a chance variable such that $2k$ different moments of Y are equal to the corresponding moments of X , then $P(Y < x)$ is identically equal to $P(X < x)$.*

From Propositions 3 and 5 we get the

PROPOSITION 7. *For each system of moments M_{i_1}, \dots, M_{i_j} there exists at most one chance variable which is characteristic relative to M_{i_1}, \dots, M_{i_j} .*

We shall now prove the

PROPOSITION 8. *If the chance variable X is characteristic relative to M_{i_1}, \dots, M_{i_j} , and M'_r is the r th moment of X , where $r > i_1, \dots, i_j$, then $M_{i_1}, \dots, M_{i_j}, M_r$ for $M_r < M'_r$ cannot be realized as moments of the orders i_1, \dots, i_j, r .*

Let us suppose that there exists a chance variable Y with the moments $M_{i_1}, \dots, M_{i_j}, M_r$ where $M_r < M'_r$. We shall deduce a contradiction from this assumption. We can assume that Y is an arithmetic chance variable, because according to a well known theorem a finite system of moments can always be realized by an arithmetic chance variable. On account of Proposition 5, $D(x) = P(Y < x) - P(X < x)$ must have at least j changes in sign. Since X is a characteristic chance variable, the number of changes in sign of $D(x)$ cannot exceed j ; hence the number of changes in sign must be equal to j . It is easy to see that the number of changes in sign can be equal to j only if the greatest value x' for which $P(Y = x') > 0$ is greater than the greatest value x'' for which $P(X = x'') > 0$. We denote by Y_d the arithmetic chance variable defined as follows:

$$P(Y_d = d) = \frac{M'_r - M_r}{d^r}, \quad P(Y_d = x') = P(Y = x') - \frac{M'_r - M_r}{d^r},$$

$$P(Y_d = x) = P(Y = x), \quad \text{for } x \neq d, x',$$

where $d > x'$ and $P(Y = x') > (M'_r - M_r)/d^r$. The differences between the moments (of the orders i_1, \dots, i_j, r) of X and the corresponding moments of Y_d become arbitrarily small if we choose d sufficiently large. It is obvious that $P(X < x) - P(Y_d < x)$ has always the same sign as $P(X < x) - P(Y < x)$. Since the number of changes in sign of $D(x)$ is equal to j , a polynomial $P(x) = a_1 x^{i_1} + \dots + a_j x^{i_j} + a_r x^r$ can be given such that $P'(x) = i_1 a_1 x^{i_1-1} + \dots + i_j a_j x^{i_j-1} + r a_r x^{r-1}$ has always the same sign as that of $P(X < x) - P(Y < x)$ and therefore has also the same sign as that of $P(X < x) - P(Y_d < x)$ for any d . Since $P(Y_d < x) = P(Y < x)$ for any $x < x'$ and since $P(Y < x) - P(X < x)$ is not equal to zero for any $x < x'$, the integral

$$\int_0^{\infty} P'(x)[P(Y_d < x) - P(X < x)]dx$$

cannot converge towards zero if $d \rightarrow \infty$. But on the other hand the moments of the order i_1, \dots, i_j, r of Y_d converge towards the corresponding moments of X if $d \rightarrow \infty$ and therefore, as can easily be shown, the above integral must converge towards zero. Hence we get a contradiction and our proposition is proved.

DEFINITION 7. A sequence $\{X_i\}$ of chance variables is said to be convergent towards the chance variable X , in symbols $\lim_{i \rightarrow \infty} X_i = X$, if $\{P(X_i < x)\}$ ($i = 1, 2, \dots$, ad inf.) converges uniformly towards $P(X < x)$ in any closed set of values of x which does not contain any point of discontinuity of $P(X < x)$.

In the following development we shall understand by "X is equal to Y," in symbols $X = Y$, that $P(X < x)$ is identically equal to $P(Y < x)$.

For any integer r we shall denote the r th moment of a chance variable X also by $M_r(X)$.

DEFINITION 8. A chance variable X_α defined for any point α of a domain D is said to be a continuous function of α in D , if for any α in D and for any sequence of points $\{\alpha_i\}$ in D which converges towards α , $\lim_{i \rightarrow \infty} X_{\alpha_i} = X_\alpha$.

Now we shall prove

PROPOSITION 9. If $\{X_i\}$ ($i = 1, 2, \dots$, ad inf.) denotes a sequence of arithmetic chance variables of degree less than a certain integer n which converges towards the chance variable X , and if for a certain positive integer r , $\{M_r(X_i)\}$ ($i = 1, 2, \dots$, ad inf.) is bounded, then $\lim_{i \rightarrow \infty} M_s(X_i) = M_s(X)$ for any positive integer $s < r$.

It is obvious that X is an arithmetic chance variable of degree less than n . Denote by $\epsilon_s(X_i, t)$ the Stieltjes integral $\int_t^\infty x^s dP(X_i < x)$ where $t > 0$. It is obvious that for any positive value t for which $P(X \geq t) = 0$

$$\lim_{i \rightarrow \infty} [M_k(X_i) - \epsilon_k(X_i, t)] = M_k(X), \quad k = 1, 2, \dots, \text{ad inf.}$$

Suppose that $\{M_r(X_i)\}$ is bounded for a certain r . Since $\epsilon_r(X_i, t) \leq M_r(X_i)$ ($i = 1, \dots$, ad inf.), $\{\epsilon_r(X_i, t)\}$ must also be bounded. That is to say, there exists a positive value N such that $\epsilon_r(X_i, t) < N$ for any integer i and for any positive value t . Hence $\epsilon_s(X_i, t) < N/t$ for $s = 1, 2, \dots, r-1$. Let us now suppose that for a certain $s < r$, $M_s(X_i)$ does not converge towards $M_s(X)$. Then a subsequence $\{X_{i_j}\}$ ($j = 1, \dots$, ad inf.) can be given such that $M_s(X_{i_j})$ converges with increasing j towards a value $M'_s \neq M_s(X)$. We choose a value t for which $P(X < t) = 1$ and $N/t < |M'_s - M_s(X)|/2$. It is obvious that for

this t , $M_s(X_i) - \epsilon_s(X_i, t)$ cannot converge towards $M_s(X)$. Hence we have a contradiction and the assumption that $M_s(X_i)$ does not converge towards $M_s(X)$ is proved to be an absurdity.

PROPOSITION 10. *If $\{X_i\}$ ($i=1, 2, \dots$, ad inf.) denotes a sequence of arithmetic chance variables of degree less than or equal to k , and if there exists an integer r for which $\{M_r(X_i)\}$ ($i=1, \dots$, ad inf.) is bounded, then there exists a subsequence $\{X_{i_j}\}$ ($j=1, \dots$, ad inf.) which is convergent.*

Since X_i is of degree less than or equal to k , there exists a subsequence $\{X'_i\}$ of $\{X_i\}$ such that the number of different values with positive probability is the same for each element of the sequence $\{X'_i\}$ ($i=1, \dots$, ad inf.). Denote this number by s . Denote by $\alpha_{i,1} < \dots < \alpha_{i,s}$ the values for which $P(X'_i = \alpha_{i,m}) > 0$ ($m=1, \dots, s$). It is obvious that there exists a subsequence $\{X'_{i_j}\}$ ($j=1, \dots$, ad inf.) of the sequence $\{X'_i\}$ such that $\lim P(X'_{i_j} = \alpha_{i_j,m})$ exists for $m=1, \dots, s$ and the sequence $\{\alpha_{i_j,m}\}$ ($j=1, \dots$, ad inf.) converges for each $m \leq s$ towards a finite value or towards infinity. Since $\{M_r(X_i)\}$ ($i=1, \dots$, ad inf.) is bounded, $\lim_{j \rightarrow \infty} P(X'_{i_j} = \alpha_{i_j,m}) = 0$ for all m for which $\lim_{j \rightarrow \infty} \alpha_{i_j,m} = \infty$. Since $\alpha_{i,1} < \alpha_{i,2} < \dots < \alpha_{i,s}$, we have $\lim_{j \rightarrow \infty} \alpha_{i_j,m} = \infty$ if $\lim_{j \rightarrow \infty} \alpha_{i_j,m-1} = \infty$. Hence there exists an integer $m' \leq s$ such that for any integer $m > m'$ and less than or equal to s , $\lim_{j \rightarrow \infty} \alpha_{i_j,m} = \infty$ and for any integer $m \leq m'$, $\lim_{j \rightarrow \infty} \alpha_{i_j,m}$ is finite. Denote $\lim_{j \rightarrow \infty} \alpha_{i_j,m}$ by α_m and $\lim_{j \rightarrow \infty} P(X'_{i_j} = \alpha_{i_j,m})$ by p_m for any $m \leq m'$. It is obvious that $\sum_{m=1}^{m'} p_m = 1$ and $\{X'_{i_j}\}$ converges towards the arithmetic chance variable X defined as follows: $P(X = \alpha_m) = p_m$ for $m \leq m'$ and $P(X = \alpha) = 0$ for any $\alpha \neq \alpha_1, \dots, \alpha_{m'}$. Hence Proposition 10 is proved.

PROPOSITION 11. *Denote by $\{X_i\}$ ($i=1, \dots$, ad inf.) a sequence of arithmetic chance variables of degree less than or equal to k for which $\{M_r(X_i)\}$ is bounded for a certain integer r . If $\{X_i\}$ does not converge towards the chance variable X , then there exists a convergent subsequence $\{X_{i_j}\}$ such that $\lim_{j \rightarrow \infty} X_{i_j} = Y \neq X$.*

Since $\{X_i\}$ does not converge towards X , there exists a positive ϵ , a sequence of numbers $\{\alpha_i\}$ contained in a closed set which does not contain any discontinuity point of $P(X < x)$, and a subsequence $\{X'_i\}$ of $\{X_i\}$ such that $|P(X'_i < \alpha_i) - P(X < \alpha_i)| > \epsilon$ for $i=1, \dots$, ad inf. Hence no subsequence of the sequence $\{X'_i\}$ can converge towards X . On account of Proposition 10 there exists a convergent subsequence $\{X'_{i''}\}$ of the sequence $\{X'_i\}$. Hence $\lim X'_{i''}$ must be different from X and our proposition is proved.

PROPOSITION 12. *Denote by $\{X_n\}$ ($n=1, \dots$, ad inf.) a sequence of arithmetic chance variables of the degree less than or equal to $(j+1)/2$, where j denotes*

a nonnegative integer. Denote further by X an arithmetic chance variable of degree less than or equal to $(j+1)/2$ for which $M_{i_1}(X), \dots, M_{i_j}(X)$ are finite and $i_1 < i_2 < \dots < i_j$ denote certain integers. If $\lim_{n \rightarrow \infty} M_{i_\nu}(X_n) = M_{i_\nu}(X)$ ($\nu = 1, \dots, j$), then $\lim X_n = X$.

Let us suppose that $\{X_n\}$ does not converge towards X but $\lim M_{i_\nu}(X_n) = M_{i_\nu}(X)$ ($\nu = 1, \dots, j$). According to Proposition 11 there exists a subsequence $\{X_{n_m}\}$ ($m = 1, \dots, \text{ad inf.}$) such that $\lim_{m \rightarrow \infty} X_{n_m} = X' \neq X$. It is obvious that X' is of degree less than or equal to $(j+1)/2$. Consider now the case that there exists an integer $r > i_j$ such that $\{M_r(X_{n_m})\}$ ($m = 1, \dots, \text{ad inf.}$) is bounded. Then we have on account of Proposition 9, $M_{i_\nu}(X) = \lim M_{i_\nu}(X_{n_m}) = M_{i_\nu}(X')$. From Proposition 5 it follows that $D(x) = P(X < x) - P(X' < x)$ must have at least j changes in sign. But this is not possible since on account of Proposition 3, $D(x)$ cannot have more than $2(j+1)/2 - 2 = j - 1$ changes in sign. Hence for any integer $r > i_j$, $\{M_r(X_{n_m})\}$ is not bounded. Hence there exists a subsequence $\{X'_{n_m}\}$ of $\{X_{n_m}\}$ such that $\lim_{m \rightarrow \infty} M_r(X'_{n_m}) = \infty$. Denote by α_m the greatest value for which $P(X'_{n_m} = \alpha_m) > 0$. Obviously $\lim \alpha_m = \infty$. Since $\lim X'_{n_m} = X'$, $\lim P(X'_{n_m} = \alpha_m)$ must be equal to zero. From this fact it follows easily that the degree of X' must be less than or equal to $(j+1)/2 - 1 = (j-1)/2$. From Proposition 9 we get that

$$M_{i_\nu}(X') = \lim M_{i_\nu}(X'_{n_m}) = M_{i_\nu}(X), \quad \nu = 1, \dots, j-1.$$

Hence according to Proposition 5, $D(x) = P(X < x) - P(X' < x)$ must have at least $j-1$ changes in sign. But this is not possible, because the degree of X' is less than or equal to $(j-1)/2$ and therefore on account of Proposition 2 the number of changes in sign of $D(x)$ is less than or equal to $2(j-1)/2 - 1 = j-2$. Hence we obtain a contradiction and our assumption that $\{X_n\}$ does not converge towards X is proved to be an absurdity.

PROPOSITION 13. Denote by M_{i_1}, \dots, M_{i_j} the moments of the orders $i_1 < i_2 < \dots < i_j$ of a certain chance variable X . There exists always a chance variable X' which is characteristic relative to M_{i_1}, \dots, M_{i_j} .

We shall prove this proposition by mathematical induction. Proposition 13 is obviously true for $j=1$. We shall suppose that 13 is true for any integer $r \leq k$. That is to say, we shall make the

ASSUMPTION A_k . Denote by M_{i_1}, \dots, M_{i_r} the moments of the orders $i_1 < \dots < i_r$ of a certain chance variable X , where $r \leq k$. There exists a chance variable X' which is characteristic relative to M_{i_1}, \dots, M_{i_r} .

In order to prove A_{k+1} , we shall first prove by means of A_k the

LEMMA B_k . If the chance variable which is characteristic relative to the moments M_{i_1}, \dots, M_{i_r} ($r \leq k$) is not degenerate, then there exists a positive δ such that any r -tuple $M'_{i_1}, \dots, M'_{i_r}$ can be realized as moments for which

$$|M_{i_1} - M'_{i_1}| < \delta, \dots, |M_{i_{r-1}} - M'_{i_{r-1}}| < \delta$$

and $M'_{i_r} > M_{i_r} - \delta$.

We shall say that an n -tuple y_1, \dots, y_n lies in the ϵ -neighborhood of the n -tuple x_1, \dots, x_n if $|x_1 - y_1| < \epsilon, \dots, |x_n - y_n| < \epsilon$.

B_k is obviously true for $r=1$. We shall prove B_k for r by assuming that it is true for $r-1$. Denote by X the chance variable which is characteristic relative to M_{i_1}, \dots, M_{i_r} and suppose that X is not degenerate. That is to say, the degree of X is equal to $(r+1)/2$. According to A_k there exists a chance variable Y which is characteristic relative to $M_{i_1}, \dots, M_{i_{r-1}}$. The chance variable Y is also not degenerate. In fact, if Y were degenerate, that is to say, if the degree of Y were less than or equal to $(r-1)/2$, then according to Proposition 6, $P(X < x)$ would be identically equal to $P(Y < x)$ and therefore also X would be degenerate, in contradiction to our assumption. Hence the degree of Y is equal to $r/2$. From Propositions 2 and 5 it follows that $M_{i_r}(Y) \neq M_{i_r}(X)$. Hence on account of Proposition 8, $M_{i_r}(Y) < M_{i_r}(X)$. Since B_k is assumed to be true for $r-1$, there exists a positive ϵ such that any $(r-1)$ -tuple $M'_{i_1}, \dots, M'_{i_{r-1}}$ in the ϵ -neighborhood of $M_{i_1}, \dots, M_{i_{r-1}}$ can be realized as moments. Hence according to A_k , for each point $M' = M'_{i_1}, \dots, M'_{i_{r-1}}$ of the ϵ -neighborhood of $M = M_{i_1}, \dots, M_{i_{r-1}}$, a chance variable (and only one) exists which is characteristic relative to M' . Denote by $X(M')$ the chance variable which is characteristic relative to M' . From Proposition 12 it follows that $X(M')$ is a continuous function of M' in the ϵ -neighborhood of M . For each point $M' = M'_{i_1}, \dots, M'_{i_{r-1}}$ of the ϵ -neighborhood of M the degree of $X(M')$ must be equal to $r/2$. In fact, if $X(M')$ were of degree less than $r/2$, then $X(M')$ would be characteristic also relative to $M'_{i_1}, \dots, M'_{i_{r-2}}$ and therefore on account of Proposition 8 not every point of a neighborhood of M' could be realized, in contradiction to the statement that every point of the ϵ -neighborhood of M can be realized. Hence the degree of $X(M')$ is equal to $r/2$ for any point M' of the ϵ -neighborhood of M . From this fact it follows easily that for any integer n the n th moment of $X(M')$ is a continuous function of M' in the ϵ -neighborhood of M . Since $X(M) = Y$ and $M_{i_r}(Y) < M_{i_r}(X)$, there exists a positive value $\delta < \epsilon$ such that for any point M' of the δ -neighborhood of M , the i_r th moment of $X(M')$ is less than $M_{i_r}(X) - \delta$. Consider a certain point $M' = M'_{i_1}, \dots, M'_{i_{r-1}}$ of the δ -neighborhood of M and the $(r-1)$ -tuple $M_{i_1}(d, \eta), \dots, M_{i_{r-1}}(d, \eta)$ defined as follows:

$$M_{i_1}(d, \eta) = \frac{M'_{i_1} - d^{i_1}\eta}{1 - \eta}, \dots, M_{i_{r-1}}(d, \eta) = \frac{M'_{i_{r-1}} - d^{i_{r-1}}\eta}{1 - \eta},$$

where d and η are positive numbers such that the $(r-1)$ -tuple $M_{i_1}(d, \eta), \dots, M_{i_{r-1}}(d, \eta)$ is contained in the δ -neighborhood of M . Denote by $X(d, \eta)$ the chance variable which is characteristic relative to $M_{i_1}(d, \eta), \dots, M_{i_{r-1}}(d, \eta)$. Denote further by $Y(d, \eta)$ the arithmetic chance variable defined as follows:

$$P[Y(d, \eta) = x] = P[X(d, \eta) = x] \cdot (1 - \eta), \quad \text{for } x \neq d,$$

$$P[Y(d, \eta) = d] = P[X(d, \eta) = d] \cdot (1 - \eta) + \eta.$$

It is obvious that the i ,th moment of $Y(d, \eta)$ is equal to M'_{i_ν} ($\nu=1, \dots, r-1$). The i ,th moment of $Y(d, \eta)$ is a continuous function of d and η . For $\eta=0$, $Y(d, \eta)$ is equal to $X(M')$ and therefore the i ,th moment of $Y(d, 0)$ is less than $M_{i_\nu}(X) - \delta$. Let us now consider two sequences of positive numbers $\{d_\nu\}$ and $\{\eta_\nu\}$ ($\nu=1, \dots, \text{ad inf.}$) such that $\lim d_\nu = \infty$, $\lim \eta_\nu = 0$, $\lim d_\nu^{i_{r-1}}\eta_\nu = 0$, and $\lim d_\nu^{i_\nu}\eta_\nu = \infty$. It is obvious that $\lim_{\nu \rightarrow \infty} M_{i_n}(d_\nu, \eta_\nu) = M'_{i_n}$ for $n=1, \dots, r-1$. Hence for sufficiently large ν the $(r-1)$ -tuple $M_{i_1}(d_\nu, \eta_\nu), \dots, M_{i_{r-1}}(d_\nu, \eta_\nu)$ lies in the δ -neighborhood of M for any positive $\bar{\eta}_\nu \leq \eta_\nu$. On the other hand the i ,th moment of $Y(d_\nu, \eta_\nu)$ converges towards infinity. If α denotes an arbitrary number greater than $M_{i_\nu}(X) - \delta$, then for sufficiently large ν the i ,th moment of $Y(d_\nu, \eta_\nu)$ will be greater than α . Since the i ,th moment of $Y(d_\nu, 0)$ is less than α , there exists a number $\bar{\eta}_\nu < \eta_\nu$ such that the i ,th moment of $Y(d_\nu, \bar{\eta}_\nu)$ is equal to α . This proves the Lemma B_k .

Now we shall prove A_{k+1} by means of A_k and B_k . Denote by $M_{i_1}, \dots, M_{i_{k+1}}$ the moments of the orders $i_1 < \dots < i_{k+1}$ of a certain chance variable X . According to A_k there exists a chance variable Y which is characteristic relative to M_{i_1}, \dots, M_{i_k} . If Y is degenerate, then according to Proposition 6, X must be equal to Y and Y is therefore characteristic also relative to $M_{i_1}, \dots, M_{i_{k+1}}$. Hence in this case A_{k+1} is proved. We have to consider only the case that Y is not degenerate. Hence the degree of Y is equal to $(k+1)/2$. On account of Proposition 8, $M_{i_{k+1}}(Y) \leq M_{i_{k+1}}$. If $M_{i_{k+1}}(Y) = M_{i_{k+1}}$, then Y is characteristic relative to $M_{i_1}, \dots, M_{i_{k+1}}$ and A_{k+1} is proved. We have to deal only with the case that $M_{i_{k+1}}(Y) < M_{i_{k+1}}$. Denote by d_0 the greatest positive value for which $P(Y=d_0) > 0$. Consider the chance variable $Y_{d,\epsilon}$ which is characteristic relative to

$$M_{i_1}(d, \epsilon) = \frac{M_{i_1} - d^{i_1}\epsilon}{1 - \epsilon}, \dots, M_{i_k}(d, \epsilon) = \frac{M_{i_k} - d^{i_k}\epsilon}{1 - \epsilon},$$

where $d > d_0$. On account of B_k , $Y_{d,\epsilon}$ exists for sufficiently small ϵ . According to Proposition 12, $\lim_{\epsilon \rightarrow 0} Y_{d,\epsilon} = Y$. Hence for sufficiently small values of ϵ ,

$Y_{d,\epsilon}$ is not degenerate. From Proposition 12 and B_k it follows easily that for any given d the set of values ϵ for which $Y_{d,\epsilon}$ exists and is not degenerate is an open set. Hence there exists a smallest value $\epsilon(d)$ for which $Y_{d,\epsilon(d)}$ is degenerate or does not exist. First we shall prove

LEMMA 1. $P(Y_{d,\epsilon} = d) = 0$ for $d > d_0$ and for any ϵ for which $Y_{d,\epsilon}$ exists.

Let us suppose that there exists a value $d > d_0$ and a positive ϵ such that $Y_{d,\epsilon}$ exists and $P(Y_{d,\epsilon} = d) > 0$. Consider the chance variable $\bar{Y}_{d,\epsilon}$ defined as follows:

$$\begin{aligned} P(\bar{Y}_{d,\epsilon} = d) &= P(Y_{d,\epsilon} = d) \cdot (1 - \epsilon) + \epsilon; \\ P(\bar{Y}_{d,\epsilon} = x) &= P(Y_{d,\epsilon} = x) \cdot (1 - \epsilon), \end{aligned} \quad \text{for } x \neq d.$$

It is obvious that $M_{i_\nu}(\bar{Y}_{d,\epsilon}) = M_{i_\nu}(Y_{d,\epsilon})$ ($\nu = 1, \dots, k$) and the degree of $\bar{Y}_{d,\epsilon}$ is not greater than the degree of $Y_{d,\epsilon}$. Hence $\bar{Y}_{d,\epsilon}$ is characteristic relative to M_{i_1}, \dots, M_{i_k} . According to Proposition 7, $\bar{Y}_{d,\epsilon}$ must be equal to Y , which is not the case, since $P(\bar{Y}_{d,\epsilon} = d) > 0$ and $P(Y = d) = 0$. Hence we have a contradiction and the assumption $P(Y_{d,\epsilon} = d) > 0$ is proved to be an absurdity.

We shall now prove the

LEMMA 2. If $d > d_0$ then for each $\epsilon < \epsilon(d)$, $P(Y_{d,\epsilon} \geq d) = 0$.

In fact $Y_{d,0} = Y$ and therefore $P(Y_{d,0} \geq d) = 0$. On account of Proposition 12, $Y_{d,\epsilon}$ is a continuous function of ϵ in the half open interval $[0, \epsilon(d))$. Hence if there exists a value $\epsilon' < \epsilon(d)$ for which $P(Y_{d,\epsilon'} \geq d) > 0$, then $P(Y_{d,\epsilon} = d) > 0$ must hold for a certain value $\epsilon = \epsilon'' \leq \epsilon'$, in contradiction to Lemma 1.

LEMMA 3. For $d > d_0$, $Y_{d,\epsilon(d)}$ exists and $P(Y_{d,\epsilon(d)} \geq d) = 0$.

Denote by $\{\epsilon_n\}$ a sequence of positive numbers for which $\epsilon_n < \epsilon(d)$ and $\lim \epsilon_n = \epsilon(d)$. Consider the corresponding sequence $\{Y_{d,\epsilon_n}\}$ of chance variables. On account of Proposition 10 there exists a convergent subsequence $\{Y_{d,\epsilon'_n}\}$ of the sequence $\{Y_{d,\epsilon_n}\}$. Denote $\lim_{n \rightarrow \infty} Y_{d,\epsilon'_n}$ by Y_d . Since according to Lemma 1, $P(Y_{d,\epsilon} \geq d) = 0$, $\{M_r(Y_{d,\epsilon'_n})\}$ is bounded for any integer r . Hence we have on account of Proposition 9

$$\lim_{n \rightarrow \infty} M_r(Y_{d,\epsilon'_n}) = M_r(Y_d)$$

for any integer r . Then from

$$M_{i_\nu} = \frac{M_{i_\nu} - d^{i_\nu} \cdot \epsilon_n}{1 - \epsilon_n}, \quad \nu = 1, \dots, k; n = 1, \dots, \text{ad inf.},$$

it follows that

$$M_{i_\nu}(Y_d) = \lim \frac{M_{i_\nu} - d^{i_\nu} \cdot \epsilon'_n}{1 - \epsilon'_n} = \frac{M_{i_\nu} - d^{i_\nu} \epsilon(d)}{1 - \epsilon(d)}, \quad \nu = 1, \dots, k.$$

Since Y_d is characteristic relative to the above moments, $Y_{d,\epsilon(d)}$ exists and is equal to Y_d . From $P(Y_{d,\epsilon_n} \geq d) = 0$ and $\lim Y_{d,\epsilon_n} = Y_{d,\epsilon(d)}$ it follows that $P(Y_{d,\epsilon(d)} > d) = 0$. Since on account of Lemma 1, $P(Y_{d,\epsilon(d)} = d) = 0$, we have $P(Y_{d,\epsilon(d)} \geq d) = 0$.

Now we are able to prove

LEMMA 4. Besides $\epsilon(d)$ no other value ϵ' can be given for which $Y_{d,\epsilon'}$ exists and is degenerate provided $d > d_0$.

Let us suppose that there exists a positive $\epsilon' \neq \epsilon(d)$ for which $Y_{d,\epsilon'}$ exists and is degenerate. Consider the chance variables $\bar{Y}_{d,\epsilon'}$ and $\bar{Y}_{d,\epsilon(d)}$ defined as follows:

$$P(\bar{Y}_{d,\epsilon} = d) = \epsilon; P(\bar{Y}_{d,\epsilon} = x) = P(Y_{d,\epsilon} = x) \cdot (1 - \epsilon), \text{ for } x \neq d; \epsilon = \epsilon', \epsilon(d).$$

Since $Y_{d,\epsilon'}$ and $Y_{d,\epsilon(d)}$ are degenerate, their degrees are less than or equal to $(k+1)/2 - 1/2 = k/2$. The degree of $\bar{Y}_{d,\epsilon'}$ and that of $\bar{Y}_{d,\epsilon(d)}$ are obviously not greater than $k/2 + 1$. Hence on account of Proposition 4,

$$D(x) = P(\bar{Y}_{d,\epsilon'} < x) - P(\bar{Y}_{d,\epsilon(d)} < x)$$

has at most $2(k/2 + 1) - 3 = k - 1$ changes in sign. Since $M_{i_\nu}(\bar{Y}_{d,\epsilon'}) = M_{i_\nu}(\bar{Y}_{d,\epsilon(d)}) = M_{i_\nu}$ ($\nu = 1, \dots, k$), $D(x)$ must be identically equal to zero on account of Proposition 5. But $D(x)$ cannot be identically equal to zero since $P(\bar{Y}_{d,\epsilon'} = d) = \epsilon'$, $P(\bar{Y}_{d,\epsilon(d)} = d) = \epsilon(d)$ and $\epsilon' \neq \epsilon(d)$. Hence the assumption that there exists an $\epsilon' \neq \epsilon(d)$ for which $Y_{d,\epsilon'}$ exists and is degenerate is proved to be an absurdity.

Let us consider a sequence of numbers $\{d_n\}$ ($n = 1, \dots, \text{ad inf.}$) for which $d_n > d_0$ and $\lim d_n = d > d_0$. We shall show that $\lim \epsilon(d_n) = \epsilon(d)$ and $\lim Y_{d_n,\epsilon(d_n)} = Y_{d,\epsilon(d)}$. In order to prove $\lim Y_{d_n,\epsilon(d_n)} = Y_{d,\epsilon(d)}$, we have only to show on account of Proposition 11 that for each convergent subsequence $\{Y_{d_n',\epsilon(d_n')}\}$ of the sequence $\{Y_{d_n,\epsilon(d_n)}\}$

$$\lim Y_{d_n',\epsilon(d_n')} = Y_{d,\epsilon(d)}.$$

Denote $\lim Y_{d_n',\epsilon(d_n')}$ by Y^* . Since $P[Y_{d_n',\epsilon(d_n')} > d_n'] = 0$, $\{M_r(Y_{d_n',\epsilon(d_n')})\}$ is bounded for any r . Hence we have on account of Proposition 9

$$\lim_{n \rightarrow \infty} M_r(Y_{d_n',\epsilon(d_n')}) = M_r(Y^*)$$

for any positive integer r . Since

$$M_{i_\nu}(Y_{d_n',\epsilon(d_n')}) = \frac{M_{i_\nu} - (d_n')^{i_\nu} \cdot \epsilon(d_n')}{1 - \epsilon(d_n')}$$

converges with increasing n and since $\lim (d_n')^{i_\nu} = d^{i_\nu} > d_0^{i_\nu} > M_{i_\nu}$, the sequence

$\{\epsilon(d_n')\}$ must also converge. Denote $\lim \epsilon(d_n')$ by ϵ^* . Then Y^* is characteristic relative to

$$M_{i_1}(d, \epsilon^*), \dots, M_{i_k}(d, \epsilon^*);$$

that is to say, Y^* is equal to Y_{d, ϵ^*} . Since $Y_{d, \epsilon^*} = \lim Y_{d_n', \epsilon(d_n')}$ and $Y_{d_n', \epsilon(d_n')}$ is degenerate, Y_{d, ϵ^*} must also be degenerate. Then according to Lemma 4, $\epsilon^* = \epsilon(d)$ and therefore Y_{d, ϵ^*} is equal to $Y_{d, \epsilon(d)}$. Hence our statement that $\lim Y_{d_n, \epsilon(d_n)} = Y_{d, \epsilon(d)}$ is proved. Since according to Lemma 3, $P(Y_{d_n, \epsilon(d_n)} \geq d_n) = 0$ and therefore $M_r(Y_{d_n, \epsilon(d_n)})$ is bounded for any integer r , we have on account of Proposition 9

$$\lim_{n \rightarrow \infty} M_r(Y_{d_n, \epsilon(d_n)}) = M_r(Y_{d, \epsilon(d)}).$$

From this it follows that $\lim \epsilon(d_n) = \epsilon(d)$ and that the moments of $Y_{d, \epsilon(d)}$ are continuous functions of d for $d > d_0$.

Denote by $\bar{Y}_{d, \epsilon}$ the chance variable defined as follows:

$$P(\bar{Y}_{d, \epsilon} = d) = \epsilon; \quad P(\bar{Y}_{d, \epsilon} = x) = P(Y_{d, \epsilon} = x) \cdot (1 - \epsilon), \quad \text{for } x \neq d.$$

It is obvious that

$$M_{i_\nu}(\bar{Y}_{d, \epsilon}) = M_{i_\nu}, \quad \nu = 1, \dots, k.$$

In order to show that $\lim_{d \rightarrow \infty} M_{i_{k+1}}(\bar{Y}_{d, \epsilon(d)}) = \infty$ we have only to show that for any sequence $\{d_n\}$ for which $\lim d_n = \infty$, $d_n^{i_k} \cdot \epsilon(d_n)$ does not converge towards zero. In order to prove the latter statement, let us assume that $\lim d_n^{i_k} \epsilon(d_n) = 0$ and $\lim d_n = \infty$. It is obvious that $\lim d_n^{i_\nu} \epsilon(d_n) = 0$ for $\nu = 1, 2, \dots, k$. Hence

$$\lim_{d \rightarrow \infty} M_{i_\nu}(Y_{d, \epsilon(d)}) = M_{i_\nu}, \quad \nu = 1, \dots, k.$$

Since Y is characteristic relative to M_{i_1}, \dots, M_{i_k} , we have, on account of Proposition 12, $\lim Y_{d_n, \epsilon(d_n)} = Y$. But this is not possible since $Y_{d_n, \epsilon(d_n)}$ is degenerate and therefore $\lim Y_{d_n, \epsilon(d_n)}$ must also be degenerate and consequently cannot be equal to Y which is not degenerate. Hence we have $\lim M_{i_{k+1}}(\bar{Y}_{d_n, \epsilon(d_n)}) = \infty$.

On account of Proposition 10 there exists a sequence $\{d_n\}$ such that $d_n > d_0$, $\lim d_n = d_0$, and the sequence $\{\bar{Y}_{d_n, \epsilon(d_n)}\}$ is convergent. Denote $\lim \bar{Y}_{d_n, \epsilon(d_n)}$ by Y^* . Since $M_{i_\nu}(\bar{Y}_{d_n, \epsilon(d_n)}) = M_{i_\nu}$ ($\nu = 1, \dots, k$) and $P(\bar{Y}_{d_n, \epsilon(d_n)} > d_n) = 0$, we have, on account of Proposition 9, $M_{i_\nu}(Y^*) = M_{i_\nu}$ ($\nu = 1, \dots, k$). The degree of $Y_{d, \epsilon(d)}$ is less than or equal to $k/2$ and therefore the degree of $\bar{Y}_{d, \epsilon(d)}$ is less than or equal to $k/2 + 1$. Hence also the degree of Y^* is less than or equal to $k/2 + 1$. Now we shall show that $P(Y^* = d_0) > 0$. Let us assume that $P(Y^* = d_0) = 0$. Then $\lim \epsilon(d_n)$ must be equal to zero. Hence $\lim M_{i_\nu}(Y_{d_n, \epsilon(d_n)})$

$=M_{i_\nu}$ ($\nu=1, \dots, k$), and then on account of Proposition 12, $Y_{d_n, \epsilon(d_n)}$ must converge towards Y which cannot be the case since $Y_{d_n, \epsilon(d_n)}$ is degenerate and Y is not degenerate. Hence $P(Y^*=d_0) > 0$ is proved. Since $P(Y=d_0) > 0$, $P(Y^*=d_0) > 0$, we get from Proposition 4 that the number of changes in sign of $D(x) = P(Y^* < x) - P(Y < x)$ is less than or equal to $2(k/2+1)-3 = k-1$. Since $M_{i_\nu}(Y^*) = M_{i_\nu}(Y)$ ($\nu=1, \dots, k$), on account of Proposition 5, $D(x)$ must be identically equal to zero, that is to say, $Y^* = Y$. Hence

$$\lim M_{i_{k+1}}(\bar{Y}_{d_n, \epsilon(d_n)}) = M_{i_{k+1}}(Y^*) = M_{i_{k+1}}(Y) < M_{i_{k+1}}.$$

Since $\lim_{d \rightarrow \infty} M_{i_{k+1}}(\bar{Y}_{d, \epsilon(d)}) = \infty$ and $M_{i_{k+1}}(\bar{Y}_{d, \epsilon(d)})$ is a continuous function of d , there exists a value d' such that $M_{i_{k+1}}(\bar{Y}_{d', \epsilon(d')}) = M_{i_{k+1}}$. The degree of $\bar{Y}_{d', \epsilon(d')}$ is less than or equal to $k/2+1$, and therefore $\bar{Y}_{d', \epsilon(d')}$ is characteristic relative to $M_{i_1}, \dots, M_{i_{k+1}}$. This proves A_{k+1} , and therefore Proposition 13 is also proved.

Since Proposition 13 is proved, B_k is also proved for any positive integer k . Hence we can formulate

PROPOSITION 14. *If the chance variable which is characteristic relative to the moments M_{i_1}, \dots, M_{i_k} is not degenerate, then there exists a positive δ such that any k -tuple $M'_{i_1}, \dots, M'_{i_k}$ in the δ -neighborhood of the k -tuple M_{i_1}, \dots, M_{i_k} can be realized as moments of the orders i_1, \dots, i_k .*

4. Solution of Problem 1. Denote by M_{i_1}, \dots, M_{i_k} the moments of the order $i_1 < i_2 < \dots < i_k$ of a certain chance variable X . Denote by X' the characteristic chance variable relative to M_{i_1}, \dots, M_{i_k} . If X' is degenerate, then according to Proposition 6 no chance variable $Y \neq X'$ exists for which $M_{i_\nu}(Y) = M_{i_\nu}(X')$ ($\nu=1, \dots, k$). Hence the sharp lower and the sharp upper limits of $P(X < d)$ are equal to $P(X' < d)$ and our problem is solved. Throughout the following development we shall suppose that X' is not degenerate.

Consider the k -tuple of values

$$M_{i_\nu}(d, \lambda) = \frac{M_{i_\nu} - d^{i_\nu} \lambda}{1 - \lambda}, \quad \nu = 1, \dots, k,$$

where $d > 0$, $0 \leq \lambda < 1$. According to Proposition 14 the k -tuple $M_{i_1}(d, \lambda), \dots, M_{i_k}(d, \lambda)$ can be realized as moments for sufficiently small values of λ . Denote by $Y(d, \lambda)$ the characteristic chance variable relative to the moments $M_{i_1}(d, \lambda), \dots, M_{i_k}(d, \lambda)$. Denote further by $\bar{Y}(d, \lambda)$ the arithmetic chance variable defined as follows:

$$\begin{aligned} P[\bar{Y}(d, \lambda) = d] &= P[Y(d, \lambda) = d](1 - \lambda) + \lambda, \\ P[\bar{Y}(d, \lambda) = x] &= P[Y(d, \lambda) = x](1 - \lambda), \end{aligned} \quad \text{for } x \neq d.$$

It is obvious that

$$M_{i_\nu}[\bar{Y}(d, \lambda)] = M_{i_\nu}, \quad \nu = 1, \dots, k.$$

From Proposition 14 it follows that for any given $d > 0$ the set Ω of values of λ for which the characteristic chance variable relative to the moments $M_{i_1}(d, \lambda), \dots, M_{i_k}(d, \lambda)$ exists and is not degenerate is an open set. Denote by λ_d the smallest positive value not belonging to Ω .

As is well known, $M_r^{s/r} \leq M_s$ for any integer $r < s$, and the equality sign holds only if the chance variable is of the degree less than or equal to 1. Since for $\lambda < \lambda_d$ the characteristic chance variable $Y(d, \lambda)$ is not degenerate, we have

$$[M_{i_\mu}(d, \lambda)]^{i_\nu/i_\mu} < M_{i_\nu}(d, \lambda), \quad \text{for } \mu < \nu, \mu < k, \nu \leq k.$$

From these inequalities and from the fact that negative moments are not possible, it follows easily that if $k \geq 3$, $\lambda_d < 1$ for any positive value d . If $k = 2$, λ_d can be equal to 1 only if $d = M_{i_1}^{1/i_1}$.

Now we shall prove

PROPOSITION 15. Denote by $\{\lambda_n\}$ ($n = 1, 2, \dots$, ad inf.) a sequence of positive values such that $\lambda_n < \lambda_d$ and $\lim \lambda_n = \lambda_d$. Then $\lim Y(d, \lambda_n)$ exists and is equal to the chance variable Y_d which is characteristic relative to the $k-1$ moments $M_{i_1}(d, \lambda_d), \dots, M_{i_{k-1}}(d, \lambda_d)$. If Y_d is not degenerate, then Y_d is characteristic also relative to the k moments $M_{i_1}(d, \lambda_d), \dots, M_{i_k}(d, \lambda_d)$.

According to Proposition 10 there exists a convergent subsequence of the sequence $\{Y(d, \lambda_n)\}$ ($n = 1, \dots$, ad inf.). Denote by $\{Y(d, \lambda'_n)\}$ a convergent subsequence of $\{Y(d, \lambda_n)\}$ and denote $\lim Y(d, \lambda'_n)$ by Y^* . If $Y(d, \lambda_d)$ exists, then according to Proposition 12, Y^* must be equal to $Y(d, \lambda_d)$. Since $Y(d, \lambda_d)$ is degenerate, $Y(d, \lambda_d)$ is characteristic also relative to the $k-1$ moments $M_{i_1}(d, \lambda_d), \dots, M_{i_{k-1}}(d, \lambda_d)$. Hence $Y^* = Y(d, \lambda_d) = Y_d$. We have now to consider the case that $Y(d, \lambda_d)$ does not exist, that is to say, the k -tuple $M_{i_1}(d, \lambda_d), \dots, M_{i_k}(d, \lambda_d)$ cannot be realized as moments. On account of Proposition 9,

$$M_{i_\nu}(Y^*) = M_{i_\nu}(d, \lambda_d), \quad \nu = 1, \dots, k-1.$$

Since the k -tuple $M_{i_1}(d, \lambda_d), \dots, M_{i_k}(d, \lambda_d)$ cannot be realized,

$$M_{i_k}(Y^*) \neq M_{i_k}(d, \lambda_d).$$

From this it follows on account of Proposition 9 that $\{M_r[Y(d, \lambda'_n)]\}$ ($n = 1, 2, \dots$, ad inf.) is not bounded for any integer $r > i_k$. Hence there exists a subsequence $\{Y(d, \lambda''_n)\}$ such that $\lim_{n \rightarrow \infty} M_r[Y(d, \lambda''_n)] = \infty$. Denote by α_n the greatest positive value for which $P[Y(d, \lambda''_n) = \alpha_n] > 0$. It is obvious that $\lim \alpha_n = \infty$ and

$$\lim P[Y(d, \lambda_n'') = \alpha_n] = 0.$$

Hence the degree of Y^* must be less than or equal to $(k+1)/2 - 1 = (k-1)/2$. Since $M_{i_\nu}(Y^*) = M_{i_\nu}(d, \lambda_d)$ ($\nu = 1, \dots, k-1$), Y^* is characteristic and degenerate relative to $M_{i_1}(d, \lambda_d), \dots, M_{i_{k-1}}(d, \lambda_d)$. That is to say, $Y^* = Y_d$ and Y_d is degenerate.

Hence we have proved that in any case the limit of a convergent subsequence of $\{Y(d, \lambda_n)\}$ is equal to Y_d . From this fact it follows on account of Proposition 11 that $\lim Y(d, \lambda_n) = Y_d$. As we have shown, $Y_d = Y(d, \lambda_d)$ if $Y(d, \lambda_d)$ exists, and Y_d is degenerate if $Y(d, \lambda_d)$ does not exist. Hence Proposition 15 is proved.

PROPOSITION 16. $P(Y_d = d) = 0$, where Y_d denotes the characteristic chance variable relative to the $k-1$ moments $M_{i_1}(d, \lambda_d), \dots, M_{i_{k-1}}(d, \lambda_d)$.

Let us suppose $P(Y_d = d) > 0$. Denote by \bar{Y}_d the chance variable defined as follows:

$$P(\bar{Y}_d = d) = P(Y_d = d) \cdot (1 - \lambda_d) + \lambda_d;$$

$$P(\bar{Y}_d = x) = P(Y_d = x) \cdot (1 - \lambda_d), \quad \text{for } x \neq d.$$

If $Y(d, \lambda_d)$ exists, then $M_{i_\nu}(Y_d) = M_{i_\nu}(d, \lambda_d)$ ($\nu = 1, \dots, k$) and therefore $M_{i_\nu}(\bar{Y}_d) = M_{i_\nu}$ ($\nu = 1, \dots, k$). The degree of \bar{Y}_d is equal to the degree of Y_d . Hence \bar{Y}_d is characteristic and degenerate relative to M_{i_1}, \dots, M_{i_k} in contradiction to our assumption that the characteristic chance variable relative to M_{i_1}, \dots, M_{i_k} is not degenerate. If $Y(d, \lambda_d)$ does not exist, then Y_d is degenerate. That is to say, the degree of Y_d is less than or equal to $(k-1)/2$. Since $M_{i_\nu}(\bar{Y}_d) = M_{i_\nu}$ ($\nu = 1, \dots, k-1$) and the degree of \bar{Y}_d is equal to the degree of Y_d , \bar{Y}_d is characteristic and degenerate relative to the $k-1$ moments $M_{i_1}, \dots, M_{i_{k-1}}$. But on account of our assumption that the characteristic chance variable relative to M_{i_1}, \dots, M_{i_k} is not degenerate, from Proposition 6 it follows that the characteristic chance variable relative to $M_{i_1}, \dots, M_{i_{k-1}}$ also cannot be degenerate. Hence we have a contradiction and the assumption that $P(Y_d = d) > 0$ is proved to be an absurdity.

PROPOSITION 17. Denote by $\{d_n\}$ and $\{\lambda_n\}$ ($n = 1, 2, \dots, ad \text{ inf.}$) two sequences of positive values such that $\lim d_n = d > 0$, $\lim \lambda_n = \lambda < \lambda_d$. Then $\lim Y(d_n, \lambda_n) = Y(d, \lambda)$.

On account of Proposition 14, $Y(d_n, \lambda_n)$ exists for almost every n . Since $\lim M_{i_\nu}[Y(d_n, \lambda_n)] = M_{i_\nu}[Y(d, \lambda)]$, we have on account of Proposition 12 that $\lim Y(d_n, \lambda_n) = Y(d, \lambda)$.

PROPOSITION 18. The sharp lower limit a_d of $P(X < d)$ is equal to $P(\bar{Y}_d < d)$, and the sharp upper limit b_d of $P(X < d)$ is equal to $P(\bar{Y}_d \leq d)$ where \bar{Y}_d denotes

the arithmetic chance variable defined as follows:

$$P(\bar{Y}_d = d) = \lambda_d; \quad P(\bar{Y}_d = x) = P(Y_d = x) \cdot (1 - \lambda_d), \quad \text{for } x \neq d.$$

We shall consider two cases.

(1) Y_d is not degenerate. Hence the degree of Y_d is equal to $k/2$. According to Proposition 15, Y_d is characteristic also relative to $M_{i_1}(d, \lambda_d), \dots, M_{i_k}(d, \lambda_d)$. Hence

$$M_{i_\nu}(\bar{Y}_d) = M_{i_\nu}, \quad \nu = 1, \dots, k.$$

Since, according to Proposition 16, $P(Y_d = d) = 0$, the degree of \bar{Y}_d is obviously equal to $k/2 + 1$. Let us suppose that there exists a chance variable X such that $M_{i_\nu}(X) = M_{i_\nu}$ ($\nu = 1, \dots, k$) and $P(X < d) < P(\bar{Y}_d < d)$. Denote by α the greatest number less than d for which $P(\bar{Y}_d = \alpha) > 0$. It is obvious that $D(x) = P(X < x) - P(\bar{Y}_d < x)$ has no change in sign in the interior of the interval $[\alpha, d]$. If $D(x)$ is identically zero in the interior of $[\alpha, d]$, then $D(x)$ has no change in sign at α . If $D(x)$ is not identically zero in the interior of $[\alpha, d]$ and if $P(X \leq d) \leq P(\bar{Y}_d \leq d)$, then $D(x)$ has no change in sign at d . Finally if $P(X \leq d) > P(\bar{Y}_d \leq d)$ and if β denotes the smallest value greater than d for which $P(\bar{Y}_d = \beta) > 0$, then $D(x)$ has no change in sign in the interior of the interval $[d, \beta]$. From this fact it follows easily that the number of changes in sign of $D(x)$ cannot exceed $2(k/2 + 1) - 3 = k - 1$. Since $M_{i_\nu}(X) = M_{i_\nu}(\bar{Y}_d)$ ($\nu = 1, \dots, k$), this is in contradiction to Proposition 5. Hence the assumption $P(X < d) < P(\bar{Y}_d < d)$ is proved to be an absurdity. Now let us assume that there exists a chance variable X such that $M_{i_\nu}(X) = M_{i_\nu}$ ($\nu = 1, \dots, k$) and $P(X < d) > P(\bar{Y}_d \leq d)$. Denote by β the smallest number greater than d for which $P(\bar{Y}_d = \beta) > 0$. It is obvious that $D(x) = P(X < x) - P(\bar{Y}_d < x)$ has no change in sign at the point d and also no change in sign in the interior of the interval $[d, \beta]$. Hence the number of changes in sign of $D(x)$ cannot exceed $2(k/2 + 1) - 3 = k - 1$. But this is in contradiction to Proposition 5, and the assumption $P(X < d) > P(\bar{Y}_d \leq d)$ therefore is proved to be an absurdity.

We now have to show that the limits $P(\bar{Y}_d < d)$ and $P(\bar{Y}_d \leq d)$ are sharp. Since $M_{i_\nu}(\bar{Y}_d) = M_{i_\nu}$ ($\nu = 1, \dots, k$), the lower limit $P(\bar{Y}_d < d)$ is evidently sharp. Denote by $\{d_n\}$ ($n = 1, 2, \dots$, ad inf.) a sequence of positive numbers for which $d_n < d$ and $\lim_{n \rightarrow \infty} d_n = d$. Denote by λ some value less than λ_d . It is obvious that $Y(d_n, \lambda)$ exists for almost every n and that on account of Proposition 12, $\lim_{n \rightarrow \infty} Y(d_n, \lambda) = Y(d, \lambda)$. Since $P(Y_d = d) = 0$ the function $P(Y_d < x)$ is constant in the neighborhood of $x = d$. Then from $\lim_{\lambda \rightarrow \lambda_d} Y(d, \lambda) = Y_d$ it follows that there exists a positive η such that

$$\lim_{\lambda \rightarrow \lambda_d} P[Y(d, \lambda) < d - \eta] = P(Y_d < d).$$

Hence to an arbitrarily small positive ϵ a value $\lambda_\epsilon < \lambda_d$ can be given such that

$$P[Y(d, \lambda) < d - \eta] > P(Y_d < d) - \epsilon$$

for any λ greater than λ_ϵ and smaller than λ_d . Since $\lim_{n \rightarrow \infty} Y(d_n, \lambda) = Y(d, \lambda)$,

$$(a) \quad P[Y(d_n, \lambda) < d] > P(Y_d < d) - 2\epsilon$$

for almost every n . On account of $d_n < d$,

$$(b) \quad P[\bar{Y}(d_n, \lambda) < d] = (1 - \lambda)P[Y(d_n, \lambda) < d] + \lambda,$$

and on account of $P(Y_d = d) = 0$,

$$(c) \quad P(\bar{Y}_d \leq d) = (1 - \lambda_d)P(Y_d < d) + \lambda_d.$$

From (a), (b), and (c), it follows that if we choose λ sufficiently near to λ_d , we have

$$P[\bar{Y}(d_n, \lambda) < d] > P(\bar{Y}_d \leq d) - 3\epsilon.$$

Since $M_{i_\nu}[\bar{Y}(d_n, \lambda)] = M_{i_\nu}$ ($\nu = 1, \dots, k$) and since ϵ can be chosen arbitrarily small, the upper limit $P(\bar{Y}_d \leq d)$ is proved to be sharp. Hence Proposition 18 is proved if Y_d is not degenerate.

(2) Y_d is degenerate. Denote the degree of Y_d by $k'/2$ where k' denotes a positive integer. It is obvious that $k' \leq k - 1$. Since the characteristic chance variable relative to M_{i_1}, \dots, M_{i_k} is not degenerate, from Proposition 6 it follows that also the characteristic chance variable relative to $M_{i_1}, \dots, M_{i_{k'}}$ is not degenerate. Considering only the moments of the orders $i_1, \dots, i_{k'}$ we have case (1) since Y_d is obviously characteristic and degenerate relative to the moments $M_1(d, \lambda_d), \dots, M_{i_{k'}}(d, \lambda_d)$. Hence $P(\bar{Y}_d < d)$ is the greatest lower and $P(\bar{Y}_d \leq d)$ is the least upper bound of $P(Z < d)$ where $P(Z < d)$ is formed for all chance variables Z for which $M_{i_\nu}(Z) = M_{i_\nu}$ ($\nu = 1, \dots, k'$).

In order to show that the lower limit $P(\bar{Y}_d < d)$ is sharp consider the sequence $\{Y(d, \lambda_n)\}$ of chance variables where $\lambda_n < \lambda_d$ and $\lim \lambda_n = \lambda_d$. Since $\lim Y(d, \lambda_n) = Y_d$ and $P(Y_d = d) = 0$, we have

$$\lim_{n \rightarrow \infty} P[Y(d, \lambda_n) < d] = P(Y_d < d).$$

On account of the fact that

$$P[\bar{Y}(d, \lambda_n) < d] = (1 - \lambda_n)P[Y(d, \lambda_n) < d],$$

and that

$$P(\bar{Y}_d < d) = P(Y_d < d) \cdot (1 - \lambda_d),$$

we have

$$\lim_{n \rightarrow \infty} P[\bar{Y}(d, \lambda_n) < d] = P(\bar{Y}_d < d).$$

Since $M_{i_\nu}[\bar{Y}(d, \lambda_n)] = M_{i_\nu}$ ($\nu = 1, \dots, k$) the lower limit $P(\bar{Y}_d < d)$ is proved to be sharp. The proof of the fact that also the upper limit $P(\bar{Y}_d \leq d)$ is sharp is quite analogous to that given in case (1). Hence Proposition 18 is proved.

We can summarize our results in the following

THEOREM 1. *The moments M_{i_1}, \dots, M_{i_j} of the orders i_1, \dots, i_j of a certain chance variable X are given. If the chance variable X' which is characteristic relative to M_{i_1}, \dots, M_{i_j} is degenerate, then the sharp lower limit a_d and the sharp upper limit b_d are equal to $P(X' < d)$. If X' is not degenerate, we have to consider the chance variable Y_d which is characteristic relative to $M_{i_1}(d, \lambda_d), \dots, M_{i_{j-1}}(d, \lambda_d)$ where*

$$M_{i_\nu}(d, \lambda) = \frac{M_{i_\nu} - d^{i_\nu} \lambda}{1 - \lambda}, \quad \nu = 1, \dots, j,$$

and λ_d denotes the smallest value λ for which $M_{i_1}(d, \lambda), \dots, M_{i_j}(d, \lambda)$ cannot be realized as moments, or the characteristic chance variable relative to them is degenerate. The sharp lower limit a_d is equal to $P(\bar{Y}_d < d)$ and the sharp upper limit b_d is equal to $P(\bar{Y}_d \leq d)$, where \bar{Y}_d denotes the arithmetic chance variable defined as follows:

$$P(\bar{Y}_d = d) = P(Y_d = d) \cdot (1 - \lambda_d) + \lambda_d,$$

$$P(\bar{Y}_d = x) = P(Y_d = x) \cdot (1 - \lambda_d) \quad \text{for} \quad x \neq d.$$

5. Solution of Problem 2. Denote by M_{i_1}, \dots, M_{i_k} the moments of the orders $i_1 < i_2 < \dots < i_k$ of a certain chance variable X . Consider an integer $i_{k+1} > i_k$ and a number $M_{i_{k+1}}$. First we shall deal with the question: what conditions must be satisfied by $M_{i_{k+1}}$ in order that $M_{i_1}, \dots, M_{i_{k+1}}$ can be realized as moments of the orders i_1, \dots, i_{k+1} .

If the chance variable Y which is characteristic relative to M_{i_1}, \dots, M_{i_k} is degenerate, then on account of Proposition 6 no chance variable $Z \neq Y$ exists such that $M_{i_\nu}(Z) = M_{i_\nu}(Y)$ ($\nu = 1, \dots, k$). Hence $M_{i_1}, \dots, M_{i_{k+1}}$ can be realized if and only if $M_{i_{k+1}} = M_{i_{k+1}}(Y)$.

Let us consider the case that Y is not degenerate. Denote by $\{d_n\}$ and $\{\epsilon_n\}$ ($n = 1, \dots$, ad inf.) two sequences of positive numbers such that $\lim d_n = \infty$, $\lim d_n^{i_\nu} \cdot \epsilon_n = 0$ for $\nu \leq k$, and $\lim d_n^{i_{k+1}} \cdot \epsilon_n = \infty$. Consider the k -tuple of values

$$M_{i_\nu}(d, \epsilon) = \frac{M_{i_\nu} - d^{i_\nu} \cdot \epsilon}{1 - \epsilon}, \quad \nu = 1, \dots, k.$$

If $M_{i_1}(d, \epsilon), \dots, M_{i_k}(d, \epsilon)$ can be realized as moments of the orders i_1, \dots, i_k ,

then we shall denote by $Y(d, \epsilon)$ the characteristic chance variable relative to these moments, and by $\bar{Y}(d, \epsilon)$ the arithmetic chance variable defined as follows:

$$\begin{aligned} P[\bar{Y}(d, \epsilon) = d] &= P[Y(d, \epsilon) = d](1 - \epsilon) + \epsilon, \\ P[\bar{Y}(d, \epsilon) = x] &= P[Y(d, \epsilon) = x](1 - \epsilon), \quad \text{for } x \neq d. \end{aligned}$$

It is obvious that

$$M_{i_\nu}[\bar{Y}(d, \epsilon)] = M_{i_\nu}, \quad \nu = 1, \dots, k.$$

Since $\lim \epsilon_n = \lim d_n \epsilon_n = 0$ ($\nu = 1, \dots, k$), from Propositions 14 and 8 it follows easily that for almost every n , $Y(d_n, \epsilon)$ exists and is not degenerate for any nonnegative value $\epsilon \leq \epsilon_n$. On account of Proposition 12, $Y(d_n, \epsilon)$ is a continuous function of ϵ in the interval $[0, \epsilon_n]$. Since $Y(d_n, \epsilon)$ is not degenerate for $0 \leq \epsilon \leq \epsilon_n$, also $M_r[Y(d_n, \epsilon)]$ is a continuous function of ϵ for any positive integer r . From this it follows that also $M_r[\bar{Y}(d_n, \epsilon)]$ is a continuous function of ϵ in the interval $[0, \epsilon_n]$. Since

$$M_{i_\nu}[\bar{Y}(d_n, \epsilon)] = M_{i_\nu} \quad (\nu = 1, \dots, k), \quad M_{i_{k+1}}[\bar{Y}(d_n, 0)] = M_{i_{k+1}}(Y),$$

we get that $M_{i_1}, \dots, M_{i_{k+1}}$ can be realized as moments if

$$M_{i_{k+1}}(Y) \leq M_{i_{k+1}} \leq M_{i_{k+1}}[\bar{Y}(d_n, \epsilon_n)].$$

Because $\lim d_n^{i_{k+1}} \epsilon_n = \infty$ we obtain easily that $\lim M_{i_{k+1}}[\bar{Y}(d_n, \epsilon_n)] = \infty$ and therefore $M_{i_1}, \dots, M_{i_{k+1}}$ can be realized as moments if $M_{i_{k+1}} \geq M_{i_{k+1}}(Y)$. From Proposition 8 it follows that this condition is also necessary. Hence we have proved

PROPOSITION 19. Denote by M_{i_1}, \dots, M_{i_k} k numbers which can be realized as moments of the orders $i_1 < i_2 < \dots < i_k$. Denote by i_{k+1} an integer greater than i_k and by $M_{i_{k+1}}$ a certain number. If the chance variable Y which is characteristic relative to M_{i_1}, \dots, M_{i_k} is degenerate, then $M_{i_1}, \dots, M_{i_{k+1}}$ can be realized as moments of the orders i_1, \dots, i_{k+1} if and only if $M_{i_{k+1}} = M_{i_{k+1}}(Y)$. If Y is not degenerate, then $M_{i_1}, \dots, M_{i_{k+1}}$ can be realized as moments if and only if $M_{i_{k+1}} \geq M_{i_{k+1}}(Y)$.

If $M_{i_{k+1}} = M_{i_{k+1}}(Y)$, the characteristic chance variable relative to $M_{i_1}, \dots, M_{i_{k+1}}$ is obviously equal to Y and therefore is degenerate. Since M_{i_1} can be realized as a moment of the order i_1 if and only if $M_{i_1} \geq 0$, we get from Proposition 19

THEOREM 2. Denote by $i_1 < i_2 < \dots < i_k$ positive integers and by M_{i_1}, \dots, M_{i_k} some numbers. The values M_{i_1}, \dots, M_{i_k} can be realized as moments of the orders i_1, \dots, i_k if and only if

$$M_{i_1} \geq 0, M_{i_2} \geq M_{i_2}(X_1), \dots, M_{i_k} \geq M_{i_k}(X_{k-1}),$$

where X_r denotes the characteristic chance variable relative to M_{i_1}, \dots, M_{i_r} ; if in one of the above relations the equality sign holds, then in all subsequent relations the equality sign must hold.

This theorem gives the solution of Problem 2, since $M_{i_r}(X_{r-1})$ is a function of $M_{i_1}, \dots, M_{i_{r-1}}$ which can be calculated.

6. Some applications of Theorems 1 and 2. Let us calculate by means of Theorem 2 the inequalities which must be satisfied by the numbers M_r, M_s, M_t if they can be realized as moments of the orders r, s, t , where $r < s < t$.

According to Theorem 2 the necessary and sufficient conditions are given by

$$(10) \quad M_r \geq 0, \quad M_s \geq M_s(X_1), \quad M_t \geq M_t(X_2),$$

where X_1 denotes the characteristic chance variable relative to M_r , and X_2 denotes the characteristic chance variable relative to M_r and M_s . The degree of X_1 is less than or equal to 1. Hence there exists only a single point a with positive probability and therefore $M_r = M_r(X_1) = a^r$. Hence $a = M_r^{1/r}$. It is obvious that

$$(11) \quad M_s(X_1) = a^s = M_r^{s/r}.$$

Let us now calculate the chance variable X_2 . The degree of X_2 is less than or equal to $3/2$. Hence only the origin and a single positive value b can have positive probability. The value of b and the probability $P(X_2 = b)$ are determined by the equations

$$M_r(X_2) = b^r P(X_2 = b) = M_r; \quad M_s(X_2) = b^s P(X_2 = b) = M_s.$$

From these equations we obtain

$$P(X_2 = b) = \frac{M_r}{(M_s/M_r)^{r/(s-r)}}, \quad b = \left(\frac{M_s}{M_r} \right)^{1/(s-r)}.$$

Hence

$$(12) \quad M_t(X_2) = b^t P(X_2 = b) = M_r \left(\frac{M_s}{M_r} \right)^{(t-r)/(s-r)}.$$

From (10), (11), and (12) we get

$$(13) \quad M_r \geq 0, \quad M_s \geq M_r^{s/r}, \quad M_t \geq M_r \left(\frac{M_s}{M_r} \right)^{(t-r)/(s-r)}.$$

If in one of the relations (13) the equality sign holds, then in all subsequent relations the equality sign must hold. These relations are necessary and suffi-

cient in order that M_r, M_s, M_t can be realized as moments of the orders r, s, t .

As an application of Theorem 1 let us calculate the sharp lower limit a_d and the sharp upper limit b_d if two moments M_r and M_s are given, where $r < s$. According to the relations (13) we have

$$M_r \geq 0, \quad M_s \geq M_r^{s/r}.$$

If $M_r = 0$ (and therefore also $M_s = 0$), or if $M_r > 0$ and $M_s = M_r^{s/r}$, the chance variable X which is characteristic relative to M_r and M_s is degenerate and we have $a_d = b_d = P(X < d)$. Since $P(X < x) = 0$ for $x \neq M_r^{1/r}$ and $P(X = M_r^{1/r}) = 1$, we have

$$\begin{aligned} a_d = b_d &= 1, & \text{for } d > M_r^{1/r}, \\ a_d = b_d &= 0, & \text{for } d \leq M_r^{1/r}. \end{aligned}$$

Now we have to consider the case that

$$(14) \quad M_r > 0, \quad M_s > M_r^{s/r}.$$

In order to calculate λ_d we have to consider the expressions:

$$M_r(d, \lambda) = \frac{M_r - d^r \lambda}{1 - \lambda}, \quad M_s(d, \lambda) = \frac{M_s - d^s \lambda}{1 - \lambda}.$$

From Theorem 2 it follows that for any λ for which $M_r(d, \lambda) > 0$ and $M_s(d, \lambda) > [M_r(d, \lambda)]^{s/r}$, $M_r(d, \lambda)$ and $M_s(d, \lambda)$ can be realized as moments of the orders r, s , and the corresponding characteristic chance variable is not degenerate. Hence either $M_r(d, \lambda_d) = 0$ or $M_s(d, \lambda_d) = [M_r(d, \lambda_d)]^{s/r}$ must hold. That is to say, λ_d is either equal to M_r/d^r or is the root of the equation

$$(15) \quad \frac{M_s - d^s \lambda}{1 - \lambda} = \left[\frac{M_r - d^r \lambda}{1 - \lambda} \right]^{s/r}.$$

We have $\lambda_d = M_r/d^r$ if and only if the smallest positive root of (15) is greater than or equal to M_r/d^r . It is easy to show that this is the case if $M_r/d^r \leq M_s/d^s$. Hence we have:

If $M_r/d^r \leq M_s/d^s$ then $\lambda_d = M_r/d^r$, and if $M_r/d^r > M_s/d^s$ then λ_d is equal to the smallest positive root of (15).

If $\lambda_d = M_r/d^r$ then the chance variable Y_d which is characteristic relative to $M_r(d, \lambda_d)$ is given as follows: $P(Y_d = 0) = 1$ and $P(Y_d = x) = 0$ for $x \neq 0$. Hence the chance variable \bar{Y}_d is given as follows:

$$\begin{aligned} P(\bar{Y}_d = 0) &= 1 - \lambda_d = 1 - M_r/d^r, \\ P(\bar{Y}_d = d) &= \lambda_d = M_r/d^r. \end{aligned}$$

Hence

$$(16) \quad a_d = P(\bar{Y}_d < d) = 1 - M_r/d^r; \quad b_d = P(\bar{Y}_d \leq d) = 1, \quad M_r/d^r \leq M_s/d^s.$$

Let us now consider the case that $M_r/d^r > M_s/d^s$. Then λ_d is the smallest positive root of (15). The chance variable Y_d which is characteristic relative to $M_r(d, y_d)$ is given as follows: $P(Y_d = \delta) = 1$ where

$$(17) \quad \delta^r = \frac{M_r - d^r \cdot \lambda_d}{1 - \lambda_d}, \quad \text{or} \quad \frac{\delta^r}{d^r} = \frac{M_r/d^r - \lambda_d}{1 - \lambda_d}.$$

The chance variable \bar{Y}_d is given as follows:

$$P(\bar{Y}_d = \delta) = P(Y_d = \delta) \cdot (1 - \lambda_d) = (1 - \lambda_d), \quad P(\bar{Y}_d = d) = \lambda_d.$$

We shall show that $\delta < d$. One can easily see that $M_r/d^r < 1$. In fact, if $M_r/d^r > 1$, then

$$\left(\frac{M_r}{d^r}\right)^{s/r} > \frac{M_r}{d^r} > \frac{M_s}{d^s}$$

and therefore $M_r^{s/r} > M_s$ which is not possible. The inequality $\delta/d < 1$ follows from (17) on account of $M_r/d^r < 1$. Hence we have

$$(18) \quad a_d = P(\bar{Y}_d < d) = P(\bar{Y}_d = \delta) = 1 - \lambda_d; \quad b_d = P(\bar{Y}_d \leq d) = 1, \\ M_r/d^r > M_s/d^s.$$

The equations (16) and (18) give the complete formulas for a_d and b_d if two moments M_r and M_s are given.

If $s = 2r$ the root λ_d of (15) is given by the expression:

$$\lambda_d = \frac{M_{2r} - M_r^2}{(d^r - M_r)^2 + (M_{2r} - M_r^2)}.$$

Hence we get

$$(19) \quad a_d = 1 - \lambda_d = 1 - \frac{M_{2r} - M_r^2}{(d^r - M_r)^2 + (M_{2r} - M_r^2)}, \\ b_d = 1, \quad M_r/d^r > M_s/d^s,$$

The sharp lower limits given in the formulas (16) and (19) are identical with the lower limits in the formulas (3) given by Cantelli.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

AN INTERPRETATION OF THE INDEX OF INERTIA OF THE DISCRIMINANT MATRICES OF A LINEAR ASSOCIATIVE ALGEBRA*

BY
R. F. RINEHART

1. **Introduction.** A famous result in the theory of algebraic equations, which was the culmination of researches of Sturm, Sylvester, Hermite, and others, is the so-called Borchardt-Jacobi Theorem, hereinafter referred to as the B. J. Theorem:† Let $f(x)=0$ be a polynomial equation of degree n with real coefficients, and let $s_i, i>0$, denote the sum of the i th powers of the roots of $f(x)=0$.

I. *The rank of the matrix*

$$T = \begin{vmatrix} n & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ \cdot & \cdot & \cdots & \cdot \\ s_{n-1} & s_n & \cdots & s_{2n-2} \end{vmatrix}$$

is equal to the number of distinct roots of $f(x)=0$.

II. *The signature of T is equal to the number of distinct real roots of $f(x)=0$.*

In the theory of linear associative algebras there exists a generalization of part I of this theorem. Let \mathfrak{A} be a linear associative algebra of order n over a field \mathfrak{K} of infinite characteristic, and let b_1, b_2, \cdots, b_n be a basis for \mathfrak{A} . Let $c_{ijk}, (i, j, k=1, \cdots, n)$, be the constants of multiplication relative to this basis. Then $b_r b_s = \sum_{i=1}^n c_{rsi} b_i, (r, s=1, \cdots, n)$. The first and second discriminant matrices of \mathfrak{A} , relative to this basis, are defined to be, respectively,

$$T_1(\mathfrak{A}) = \|t_1(b_r b_s)\| = \left\| \sum_{i=1}^n c_{rsi} t_1(b_i) \right\| = \left\| \sum_{i,j=1}^n c_{rsi} c_{ijj} \right\|,$$

$$T_2(\mathfrak{A}) = \|t_2(b_r b_s)\| = \left\| \sum_{i=1}^n c_{rsi} t_2(b_i) \right\| = \left\| \sum_{i,j=1}^n c_{rsi} c_{jii} \right\|,$$

where $t_1(b_i)$ and $t_2(b_i)$ are respectively the first and second traces of the element b_i , that is, the traces of the first and second matrices, $\|c_{isr}\|$ and $\|c_{ris}\|$,

* Presented to the Society, November 26, 1938; received by the editors March 6, 1939.

† For a complete historical account of this theorem, see the tract *Abhandlung über die Auflösung der numerischen Gleichungen* (Ostwald's Klassiker der exakten Wissenschaften, no. 143), by C. Sturm, edited by A. Loewy, Leipzig, 1904.

of the element b_i . It has been shown that $T_1(\mathfrak{A})$ and $T_2(\mathfrak{A})$ are symmetric,* and that under a transformation of basis of \mathfrak{A} , $b'_i = \sum_{j=1}^n m_{ij}b_j$, ($i=1, \dots, n$), of matrix $M = \|m_{rs}\|$, $|m_{rs}| \neq 0$, the discriminant matrices are transformed by congruence,*† namely,

$$T'_1 = MT_1M^T, \quad T'_2 = MT_2M^T, \ddagger$$

so that the ranks (and signatures, if \mathfrak{K} is an ordered field) of T_1 and T_2 are invariant under transformation of basis of \mathfrak{A} . The following theorem is well known in the theory of linear algebras:

THEOREM A. § *The nullity of $T_1(\mathfrak{A})$ [or $T_2(\mathfrak{A})$] is equal to the order of the radical of \mathfrak{A} .*

MacDuffee (cf. M1) has pointed out that the discriminant matrices of the polynomial algebra generated by an element x whose minimum equation is the polynomial equation $f(x)=0$ of degree n , relative to the basis $1, x, x^2, \dots, x^{n-1}$ become the matrix T of the B. J. Theorem. It has also been noted that, for such an algebra, Theorem A specializes precisely to part I of the B. J. Theorem,|| so that Theorem A is a direct extension of part I of the B. J. Theorem from the case of a polynomial algebra to that of an arbitrary associative algebra.

From this standpoint it is apparent that Theorem A constitutes an incomplete generalization of the B. J. Theorem. An extension of part II of the B. J. Theorem to an arbitrary algebra¶ would be desirable. Moreover, when the ground field \mathfrak{K} of the algebra \mathfrak{A} is the real field, the rank and signature of $T_1(\mathfrak{A})$ [$T_2(\mathfrak{A})$] constitute a complete set of invariants of $T_1(\mathfrak{A})$ [$T_2(\mathfrak{A})$] under transformations of basis of \mathfrak{A} . Thus, in view of Theorem A, if an interpretation of the signature (or any second invariant which is independent of the rank) of $T_1(\mathfrak{A})$ [$T_2(\mathfrak{A})$] is found, then the significance of the discriminant matrices of an algebra over the real field will be, in a sense, fully known.

It is the purpose of this paper to complete the generalization of the B. J. Theorem, and thus exhibit the significance of a complete set of invariants,

* C. C. MacDuffee, *The discriminant matrices of a linear associative algebra*, Annals of Mathematics, (2), vol. 32 (1931), pp. 60-66; hereinafter referred to as M1.

† C. C. MacDuffee, *The discriminant matrix of a semisimple algebra*, these Transactions, vol. 33 (1931), pp. 425-432; hereinafter referred to as M2. E. Noether, *Mathematische Zeitschrift*, vol. 30 (1929), p. 689.

‡ M^T denotes the transpose of M .

§ Cf. L. E. Dickson, *Algebren und ihre Zahlentheorie*, Zurich, 1927, pp. 108-110.

|| R. F. Rinehart, *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 570-576; hereinafter referred to as R1.

¶ Hereafter the term *algebra* will be understood to denote a linear associative algebra of finite order.

over the real field, of $T_1(\mathfrak{A}) [T_2(\mathfrak{A})]$. The second invariant of $T_1(\mathfrak{A}) [T_2(\mathfrak{A})]$ which seems to be most easily interpreted is μ , the number of nonnegative terms in a diagonal canonical form of $T_1(\mathfrak{A}) [T_2(\mathfrak{A})]$.* In terms of the order n , rank ρ , and signature σ of $T_1(\mathfrak{A}) [T_2(\mathfrak{A})]$, $\mu = n - (\rho - \sigma)/2$. The method of attack on the problem of interpretation is simple in motif but somewhat complicated in the details. In §2 it is shown that if \mathfrak{A} is simple, μ is equal to the number in a complete set of primitive idempotents of \mathfrak{A} , plus the order of a nilpotent subalgebra of \mathfrak{A} of maximal order. In §3 the results of §2 are extended to semisimple algebras by the obvious device of applying the classical theorem concerning the decomposition of a semisimple algebra into a direct sum of simple algebras. In §§4 and 5 the results of §3 are generalized to an arbitrary algebra by again making use of a well known structure theorem to the effect that an arbitrary algebra is the sum of its radical and semisimple algebra.† In §6 it is shown that the general theorem of §5 specializes to part II of the B. J. Theorem, when the algebra is taken to be a polynomial algebra.

2. The inertia of the discriminant matrix of a simple algebra.‡ Let \mathfrak{D} be a division algebra over the real field \mathfrak{R} . Then, as is well known, \mathfrak{D} is equivalent to one of (I) the real field \mathfrak{R} ; (II) the complex field \mathbb{C} ; (III) the algebra of real quaternions \mathbb{Q} . If we choose the customary canonical bases

$$(I) \quad 1: 1^2 = 1,$$

$$(II) \quad 1, i: 1 \cdot i = i \cdot 1 = i, i^2 = -1,$$

$$(III) \quad 1, i, j, k: 1 \cdot i = i \cdot 1 = i, 1 \cdot j = j \cdot 1 = j, 1 \cdot k = k \cdot 1 = k, 1^2 = 1, i^2 = j^2 = k^2 = -1, ij = -ji = k, jk = -kj = i, ki = -ik = j,$$

respectively, in cases (I), (II), and (III), the discriminant matrix of \mathfrak{D} assumes the respective forms

$$(I): \quad \|1\|, \quad (II): \quad \begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix}, \quad (III): \quad \begin{vmatrix} 4 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & -4 \end{vmatrix}.$$

* It is shown in §4 that the signatures (and consequently the invariants μ) of $T_1(\mathfrak{A})$ and $T_2(\mathfrak{A})$ are equal.

† Here difficulty is encountered because, while the interpretation is additive under the operation of "tacking on a radical" to a semisimple algebra, it is not easy to show that μ possesses the additive property. It seems to the writer that the fundamental Theorem 4.1 should be susceptible of a simpler proof, but such a proof was not found.

‡ MacDuffee (M2) has shown that the first and second discriminant matrices of a semisimple algebra over a field of infinite characteristic are equal relative to any given basis. Consequently, for semisimple algebras, the phrase *the discriminant matrix* is unambiguous. (The terminology *infinite characteristic* is used in lieu of the customary term *characteristic 0*. As has been noted by A. A. Albert (*Modern Higher Algebra*, University of Chicago Press, 1937), the former nomenclature seems to be more harmonious with the general definition of the characteristic in other cases.)

In each case the index of inertia μ of the discriminant matrix is unity, and no clue as to the interpretation of μ is apparent from these instances. Let us investigate the most general type of simple algebra over \mathfrak{K} .

Let \mathfrak{S} be a simple algebra over \mathfrak{K} . By Wedderburn's well known theorem, \mathfrak{S} is equivalent* to a total matrix algebra \mathfrak{M} over a division algebra \mathfrak{D} . As remarked above \mathfrak{D} must be equivalent to one of \mathfrak{K} , \mathfrak{C} , or \mathfrak{Q} . To interpret the index of inertia of $T(\mathfrak{S})$, the following theorem (which was discovered inductively) is of primary importance:

THEOREM 2.1. *Let \mathfrak{S} be a simple algebra over \mathfrak{K} of order δn^2 , where $\delta = 1, 2$, or 4 according as \mathfrak{D} is \mathfrak{K} , \mathfrak{C} , or \mathfrak{Q} . The order of a nilpotent subalgebra of \mathfrak{S} of maximal order is $\delta n(n-1)/2$.*

We note first that if $n = 1$, \mathfrak{S} is a division algebra and hence possesses no nilpotent elements, so that the order of a nilpotent subalgebra of maximal order is zero. Thus Theorem 2.1 is verified when $n = 1$.

Now let $n > 1$, and let e_{pq} , ($p, q = 1, 2, \dots, n$), be the customary basis for the total matrix algebra \mathfrak{M} ; that is, a basis having the multiplication table

$$e_{pq}e_{lm} = \delta_{ql}e_{pm}, \quad p, q, l, m = 1, \dots, n,$$

where δ_{ql} is Kronecker's delta. Let \mathfrak{V} denote the linear form module over \mathfrak{D} , a basis for which is e_{pq} , ($p = 1, \dots, n-1; q = p+1, \dots, n$). Then \mathfrak{V} is composed of all matrices of the form

$$\begin{pmatrix} 0 & d_{12} & d_{13} & \cdots & d_{1n} \\ 0 & 0 & d_{23} & \cdots & d_{2n} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & d_{n-1n} \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

where the d_{pq} are in \mathfrak{D} . It is clear that the product of any two elements of \mathfrak{V} is again in \mathfrak{V} so that \mathfrak{V} is an algebra. Furthermore, it is apparent that \mathfrak{V} is nilpotent, since the n th power of any matrix of the above form is zero.† Hence \mathfrak{V} is a nilpotent subalgebra of \mathfrak{M} . Its order is $n(n-1)/2$. Hence \mathfrak{S} has a nilpotent subalgebra of order $\delta n(n-1)/2$. A basis for this subalgebra is $d_h e_{pq}$, $h = 1, \dots, \delta$; $p = 1, \dots, n-1$; $q = p+1, \dots, n$, where the d_h are basis elements of \mathfrak{D} .

We wish to show that $\delta n(n-1)/2$ is the maximal order that a nilpotent

* Two algebras \mathfrak{A} and \mathfrak{B} will be said to be *equivalent*, if a simple ring isomorphism exists between the elements of \mathfrak{A} and those of \mathfrak{B} .

† To prove these statements one needs only the assumption of the associativity of the elements of the matrices of \mathfrak{V} .

subalgebra of \mathfrak{S} may have. For this purpose we need a lemma which we now interrupt the proof of Theorem 2.1 to establish.

LEMMA 1. *If a set of n matrices of order n of the type*

$$M_h = \sum_{l=1}^n d_{hl} e_{hl}, \quad h = 1, \dots, n,$$

where the d_{hl} are elements of a division algebra \mathfrak{D} , is such that every linear combination of them, with coefficients in the ground field \mathfrak{K} of \mathfrak{D} , is nilpotent, then one of the M_h is zero.

The proof will be made by mathematical induction on n . If $n=1$, $M_1 = (d_{11})$, where d_{11} is in \mathfrak{D} . The hypothesis that M_1 is nilpotent implies that $d_{11}=0$. Thus Lemma 1 holds for $n=1$.

Now assume the lemma to be true for order $n-1$, and consider the case of order n . Then

$$M_1 = \begin{vmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ 0 & 0 & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & 0 \end{vmatrix}, M_2 = \begin{vmatrix} 0 & 0 & \cdots & 0 \\ d_{21} & d_{22} & \cdots & d_{2n} \\ . & . & \cdots & . \\ 0 & 0 & \cdots & 0 \end{vmatrix}, \dots, M_n = \begin{vmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ . & . & \cdots & . \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{vmatrix}.$$

Consider the matrices of the form

$$\overline{M}_1 = c_2 M_2 + c_3 M_3 + \cdots + c_n M_n = \begin{vmatrix} 0 & 0 & \cdots & 0 \\ c_2 d_{21} & c_2 d_{22} & \cdots & c_2 d_{2n} \\ . & . & \cdots & . \\ c_n d_{n1} & c_n d_{n2} & \cdots & c_n d_{nn} \end{vmatrix},$$

where the c_h are arbitrary elements of \mathfrak{K} . By the hypothesis of the lemma, every such matrix must be nilpotent. This evidently implies that the submatrix of order $n-1$, which is composed of the last $n-1$ rows and columns of \overline{M}_1 , is nilpotent. Since the c_h are arbitrary, the assumption of the truth of the lemma for matrices of order $n-1$ implies that one of the rows of this submatrix consists of zeros. Hence one of the matrices M_h , say M_{h_1} , is of the form

$$M_{h_1} = \begin{vmatrix} 0 & 0 & \cdots & 0 \\ . & . & \cdots & . \\ d_{h_1 1} & 0 & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & 0 \end{vmatrix}.$$

If $d_{h_1 1}=0$, the lemma is proved for the case $n=n$. If $d_{h_1 1} \neq 0$, consider the

matrix $\overline{M}_2 = c_1 M_1 + c_3 M_3 + \cdots + c_n M_n$, where the c_h are arbitrary numbers of \mathfrak{R} . As before, the hypothesis that \overline{M}_2 is nilpotent implies that the submatrix of \overline{M}_2 of order $n-1$, which is composed of rows and columns 1, 3, 4, \dots , n of \overline{M}_2 , is nilpotent. Again, from the assumption of the induction and from the nature of the c_h , one of the M_h , say M_{h_2} , is of the form

$$M_{h_2} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & d_{h_2 2} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Furthermore, since $d_{h_1 1} \neq 0$, it follows that $h_2 \neq h_1$. As before, if $d_{h_2 2} = 0$, the lemma is proved for $n=n$. If $d_{h_2 2} \neq 0$, we proceed as in the previous instances, forming the matrix \overline{M}_3 , and find that one of the M_h say M_{h_3} , with $h_3 \neq h_1, h_2$, consists of zero elements with the possible exception of the element in the $h_3, 3$ position. In the continuation of this process we must finally arrive at an M_{h_s} which is zero. For, if this were not the case, the matrix $M = M_1 + M_2 + \cdots + M_n$ would have exactly n nonzero elements, no two of which would lie in a common row or column. As in the theory of matrices with commutative elements, such a matrix cannot be nilpotent. Indeed, it is easily seen that any power of M will again be a matrix with at most one nonzero element in each row and column. Each such element is a product of nonzero elements of \mathfrak{D} , and since \mathfrak{D} is a division algebra, no such product is zero. Thus M is not nilpotent; but this contradicts the hypothesis of the lemma. Therefore some M_h is zero, and the lemma is proved.

We return now to Theorem 2.1. We shall make the proof that $\delta n(n-1)/2$ is the maximum possible order for a nilpotent subalgebra of \mathfrak{S} , by mathematical induction on n . If $n=1$, \mathfrak{S} has no nilpotent subalgebra, and the formula $\delta n(n-1)/2$ holds.

Assume the formula holds for a total matrix algebra of order $n-1$ over \mathfrak{D} and consider the case $n=n$. Suppose that \mathfrak{S} contains a nilpotent subalgebra \mathfrak{Q}' of order $t > \delta n(n-1)/2$. We shall show that this assumption leads to a contradiction. Let l_1, l_2, \dots, l_t be a basis for \mathfrak{Q}' . Since $d_h e_{pq}$, ($h=1, \dots, \delta$; $p, q=1, \dots, n$), constitute a basis for \mathfrak{S} , each l_g is expressible as

$$(2.1) \quad l_g = \sum_{h,p,q} c_{ghpq} d_h e_{pq}, \quad g = 1, 2, \dots, t,$$

where the c_{ghpq} are in \mathfrak{R} . Now $t > \delta n(n-1)/2 \geq \delta(n-1)$. It is therefore possible to eliminate from the right-hand side of (2.1) all terms involving e_{pr} , $p \neq r$, for a fixed index r , by forming a proper linear combination, with real coeffi-

cients, of the l_g , ($g=1, \dots, t$). Such an element of \mathfrak{L}' is of the form

$$(2.2) \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} & \cdots & a_{rn} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix},$$

where the a_{pq} are in \mathfrak{D} . This matrix is clearly not nilpotent, unless $a_{rr}=0$; therefore, when the e_{pr} , $p \neq r$, for a fixed index r , are eliminated from (2.1), e_{rr} is eliminated also.

Since the l_g are linearly independent over \mathfrak{K} , and since $t > \delta n(n-1)/2$, it follows from the theory of linear dependence that the number of linearly independent elements of \mathfrak{L}' , of the form (2.2), for a fixed r , is greater than

$$\frac{1}{2}\delta n(n-1) - \delta(n-1) = \frac{1}{2}\delta(n-1)(n-2).$$

Consider the set of all elements of \mathfrak{L}' of the form (2.2) for a fixed r . Every integral rational function of these elements with real coefficients is nilpotent. However, in any such rational integral function, the elements (of the resulting matrix) in the positions h, m , $h \neq r$ and $m \neq r$, are determined completely by the elements of the matrices (2.2) in rows other than the r th and columns other than the r th. In other words the elements of the matrices (2.2) in the r th row or r th column have no effect on rows or columns other than the r th row or column. Therefore the submatrices obtained from (2.2) by deleting row r and column r constitute a nilpotent subalgebra of a total matrix algebra of order $(n-1)$ over \mathfrak{D} . By the assumption of the induction that the theorem holds for total matrix algebras of order $(n-1)$, there can be at most $\delta(n-1)(n-2)/2$ linearly independent such submatrices of the set (2.2). Since the number of linearly independent matrices of the form (2.2) is greater than $\delta(n-1)(n-2)/2$, we can, by taking a linear combination of the matrices (2.2), produce a nonzero matrix of the form

$$(2.3) \quad \begin{vmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ b_{r1} & \cdots & b_{rr-1} & 0 & b_{rr+1} & \cdots & b_{rn} \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{vmatrix}.$$

This matrix belongs of course to \mathfrak{L}' .

In the above argument r was fixed but arbitrary. Hence a nonzero matrix of \mathfrak{L}' of the form (2.3) can be constructed for every r from 1 to n . Further, any linear combination of such matrices, with real coefficients, is again in \mathfrak{L}' , and is therefore nilpotent. But this contradicts Lemma 1. Hence the assumption that \mathfrak{S} , of order δn^2 , contains a nilpotent subalgebra of order greater than $\delta n(n-1)/2$, together with the assumption of the truth of the theorem for smaller values of n , leads to a contradiction. This completes the induction proof of Theorem 2.1.

We remark in passing that the nilpotent subalgebra of \mathfrak{S} of order $\delta n(n-1)/2$ is by no means unique. There are many such subalgebras. If a similarity transformation is performed on the elements of one such algebra, one obtains another such algebra, which is equivalent to the first. Whether or not any two nilpotent subalgebras of \mathfrak{S} of maximal order are equivalent is a question that the writer has not yet investigated.

We are now in a position to prove

THEOREM 2.2. *Let μ be the index of inertia of the discriminant matrix of a simple algebra \mathfrak{S} over the real field. Let ϵ be the number in a complete set of primitive idempotents of \mathfrak{S} , and let χ be the order of a nilpotent subalgebra of \mathfrak{S} of maximal order. Then $\mu = \epsilon + \chi$.*

As previously noted, \mathfrak{S} is either (I) \mathfrak{R} , (II) \mathfrak{C} , (III) \mathfrak{Q} , or (IV) a total matrix algebra of order greater than one over one of \mathfrak{R} , \mathfrak{C} , or \mathfrak{Q} . In cases (I), (II), and (III) \mathfrak{S} is a division algebra, and hence has no nilpotent elements. Furthermore, it possesses no idempotents other than the principal unit.* Hence $\chi = 0$ and $\epsilon = 1$. We have seen that if \mathfrak{S} is a division algebra, $\mu = 1$. Hence Theorem 2.2 holds in cases (I), (II), and (III).

At this point let us recall the following known results:

(a) The discriminant matrix of the direct product of two semisimple algebras \mathfrak{A} and \mathfrak{B} is (for proper choice and ordering of the basis elements) a direct product of the discriminant matrices of \mathfrak{A} and \mathfrak{B} (cf. M2).

(b) The signature of a direct product of two symmetric matrices is equal to the product of the signatures of those matrices.†

(c) The signature of the discriminant matrix of a total matrix algebra of order n^2 over the real field is n (cf. M2).

From properties (a), (b), and (c), it follows that, in case (IV), the signature, $\sigma(T(\mathfrak{S}))$, of $T(\mathfrak{S})$ is n , 0, or $-2n$, according as \mathfrak{D} is \mathfrak{R} , \mathfrak{C} , or \mathfrak{Q} . For any symmetric matrix, $\mu = (\rho + \sigma)/2$, where ρ is the rank of the matrix. Since \mathfrak{S} is simple, $T(\mathfrak{S})$ is nonsingular, and hence $\rho(T(\mathfrak{S})) = \delta n^2$. Hence according as

* Cf. L. E. Dickson, op. cit., p. 112.

† Cf. C. C. MacDuffee, *The Theory of Matrices*, Springer, Berlin, 1933, p. 83.

\mathfrak{D} is \mathfrak{R} , \mathfrak{C} , or \mathfrak{Q} , we have, respectively,

- (1) $\mu(T(\mathfrak{S})) = (n^2 + n)/2$,
- (2) $\mu(T(\mathfrak{S})) = (2n^2)/2 = n^2$,
- (3) $\mu(T(\mathfrak{S})) = (4n^2 - 2n)/2 = 2n^2 - n$.

Now the number in a complete set of primitive idempotents of a total matrix algebra over a division algebra is easily seen to be the same as the number of such idempotents of a total matrix algebra over a field, namely n . By Theorem 2.1 the order χ of a nilpotent subalgebra of \mathfrak{S} of maximal order is $\delta n(n-1)/2$, where $\delta = 1, 2$, or 4 , respectively, in cases (1), (2), and (3). Hence in the three cases we have

- (1) $\chi + \epsilon = n(n-1)/2 + n = n(n+1)/2 = \mu(T(\mathfrak{S}))$,
- (2) $\chi + \epsilon = n(n-1) + n = n^2 = \mu(T(\mathfrak{S}))$,
- (3) $\chi + \epsilon = 2n(n-1) + n = 2n^2 - n = \mu(T(\mathfrak{S}))$,

which completes the proof.

It may occur to the reader at this point that Theorem 2.2 can be proved for the more general case where the ground field is any ordered field, for instance the rational field. However, the number of primitive idempotents of an algebra is not invariant under change of ground field, so that Theorem 2.2 is not valid, in general, for an arbitrary ordered field, and in particular, is not valid, in general, for the rational field.

Theorem 2.2 can be put into the alternative form:

THEOREM 2.3. *Let \mathfrak{S} be a simple algebra over \mathfrak{R} , and let \mathfrak{B} be a subalgebra of \mathfrak{S} of minimum order which contains a complete set of primitive idempotents of \mathfrak{S} , and which has, as its radical, a nilpotent subalgebra of \mathfrak{S} of maximum order. Then the order of \mathfrak{B} is equal to $\mu(T(\mathfrak{S}))$.*

To prove this theorem it is sufficient to exhibit a \mathfrak{B} whose order is $\mu(T(\mathfrak{S}))$, since no algebra of the type of \mathfrak{B} of the theorem can have an order smaller than $\mu(T(\mathfrak{S}))$. Such an algebra is that of all matrices of the form

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdot & \cdots & a_{nn} \end{vmatrix},$$

where the a_{rr} are arbitrary real numbers, and the a_{rs} , $r < s$, are arbitrary elements of \mathfrak{D} .

3. Extension to semisimple algebras. The method of extension of the results of §2 to a semisimple algebra is fairly apparent. Let \mathfrak{A} be a semisimple

algebra over \mathfrak{K} . By the well known decomposition theorem, \mathfrak{A} is equivalent to a direct sum of simple algebras $\mathfrak{S}_1, \mathfrak{S}_2, \dots, \mathfrak{S}_\beta$. It is clear that a nilpotent subalgebra of \mathfrak{A} of maximal order will be a direct sum of such nilpotent subalgebras of the \mathfrak{S}_α . Further, a complete set of primitive idempotents of \mathfrak{A} will be composed of the complete sets of primitive idempotents of the \mathfrak{S}_α .

On the other hand, for a proper choice of basis of \mathfrak{A} , $T(\mathfrak{A})$ is a direct sum of the discriminant matrices of the \mathfrak{S}_α (cf. M2). Moreover, the rank of a direct sum of matrices is equal to the sum of the ranks of the component matrices, and the same is true of the signature when the matrices are symmetric. Hence the index of inertia of $T(\mathfrak{A})$ is equal to the sum of the indices of inertia of the $T(\mathfrak{S}_\alpha)$. This proves

THEOREM 3.1. *Let \mathfrak{A} be a semisimple algebra over \mathfrak{K} , and let ϵ be the number in a complete set of primitive idempotents of \mathfrak{A} , and χ the order of a nilpotent subalgebra of \mathfrak{A} of maximal order. Then $\mu(T(\mathfrak{A})) = \chi + \epsilon$.*

It is clear that Theorem 2.3 becomes

THEOREM 3.2. *Let \mathfrak{A} be a semisimple algebra over \mathfrak{K} , and let \mathfrak{B} be a subalgebra of \mathfrak{A} of minimum order, which contains a complete set of primitive idempotents of \mathfrak{A} , and which has, as its radical, a nilpotent subalgebra of \mathfrak{A} of maximum order. Then the order of \mathfrak{B} is equal to $\mu(T(\mathfrak{A}))$.*

Let \mathfrak{A} be an algebra of order n over a subfield \mathfrak{K} of the real field \mathfrak{R} . If the signature of $T_1(\mathfrak{A})$ is equal to n , $T_1(\mathfrak{A})$ is nonsingular, and therefore \mathfrak{A} is semisimple. Then $T_1(\mathfrak{A}) = T_2(\mathfrak{A}) = T(\mathfrak{A})$. Let \mathfrak{A}' denote the algebra \mathfrak{A} taken over the real field. Then \mathfrak{A}' is equivalent to a direct sum of simple algebras each of which has a discriminant matrix whose signature is equal to its order. From §2 the only simple algebra whose order is equal to the signature of its discriminant matrix is the real field itself. Hence \mathfrak{A}' is equivalent to a direct sum of algebras of order one, each of which is equivalent to \mathfrak{R} . Consequently \mathfrak{A}' , and therefore \mathfrak{A} , is commutative. From the theory of polynomial algebras, every semisimple algebra over a field \mathfrak{K} of infinite characteristic is equivalent to the polynomial algebra generated by a polynomial with coefficients in \mathfrak{K} and without repeated factors.* This proves

THEOREM 3.3. *Let \mathfrak{A} be an algebra of order n over a subfield \mathfrak{K} of \mathfrak{R} . If the signature of $T_1(\mathfrak{A})$ [$T_2(\mathfrak{A})$] is n , then \mathfrak{A} is equivalent to a polynomial algebra generated by a polynomial of degree n with coefficients in \mathfrak{K} and without repeated factors, and \mathfrak{A} is therefore commutative.*

4. The fundamental theorem for the extension to an arbitrary algebra.
Let \mathfrak{A} be an arbitrary non-nilpotent associative algebra over \mathfrak{K} . Then \mathfrak{A} is

* Cf. R. F. Rinehart, *Commutative algebras which are polynomial algebras*, Duke Mathematical Journal, vol. 4 (1938), p. 725; hereinafter referred to as R2.

the sum of its radical \mathfrak{Z} , and a semisimple algebra \mathfrak{A} , which is equivalent to the difference algebra $\mathfrak{A}/\mathfrak{Z}$.† If a basis for \mathfrak{A} is chosen to consist of a basis for \mathfrak{A}^* together with a basis for \mathfrak{Z} , the first and second discriminant matrices of \mathfrak{A} take the form

$$T_1(\mathfrak{A}) = \begin{vmatrix} A_1 & 0 \\ 0 & 0 \end{vmatrix}, \quad T_2(\mathfrak{A}) = \begin{vmatrix} A_2 & 0 \\ 0 & 0 \end{vmatrix},$$

where A_1 and A_2 are nonsingular square matrices, whose order is the order of \mathfrak{A}^* . The matrix A_1 [A_2] is $\|t_1(a, a_s) \|$ [$\|t_2(a, a_s) \|$], where the a_h are basis elements of \mathfrak{A}^* , and where $t_1(a, a_s)$ [$t_2(a, a_s)$] is the trace of the first [second] matrix of the element a, a_s in the representation of \mathfrak{A} by its first [second] matrices (cf. M2 and R1).‡

In working with an algebra \mathfrak{C} and a subalgebra \mathfrak{B} of \mathfrak{C} , the notation $T_1(\mathfrak{B})$ [$T_2(\mathfrak{B})$], relative to a given basis of \mathfrak{B} , is ambiguous. For $T_1(\mathfrak{B})$ [$T_2(\mathfrak{B})$] may be formed from the traces of the matrices of the elements of \mathfrak{B} in the representation of \mathfrak{B} by its first [second] matrices, or from the traces of the matrices of the elements of \mathfrak{B} in the representation of \mathfrak{C} by its first [second] matrices. To avoid this ambiguity we introduce the notation $_{\mathfrak{C}}T_1(\mathfrak{B})$ [$_{\mathfrak{C}}T_2(\mathfrak{B})$], to indicate that $T_1(\mathfrak{B})$ [$T_2(\mathfrak{B})$] is formed from the traces of the matrices in the representation of \mathfrak{C} by first [second] matrices. For $_{\mathfrak{B}}T_1(\mathfrak{B})$ [$_{\mathfrak{B}}T_2(\mathfrak{B})$] we shall write simply $T_1(\mathfrak{B})$ [$T_2(\mathfrak{B})$], when no confusion is likely to result.

In terms of this notation it is readily seen that the matrices A_1 and A_2 of the first paragraph are, respectively, $_{\mathfrak{A}}T_1(\mathfrak{A}^*)$ and $_{\mathfrak{A}}T_2(\mathfrak{A}^*)$, relative to the basis chosen for \mathfrak{A} . The ranks of $_{\mathfrak{A}}T_1(\mathfrak{A}^*)$, $_{\mathfrak{A}}T_2(\mathfrak{A}^*)$, and $T(\mathfrak{A}^*)$ are equal, for, since \mathfrak{A}^* is semisimple, $T(\mathfrak{A}^*)$ is nonsingular. As a first step in the extension of Theorem 3.1 to an arbitrary algebra we shall prove that the signatures of $_{\mathfrak{A}}T_1(\mathfrak{A}^*)$, $_{\mathfrak{A}}T_2(\mathfrak{A}^*)$, and $T(\mathfrak{A}^*)$ are likewise equal. For this purpose we need several lemmas, which we shall establish presently.

Let \mathfrak{S} be a simple algebra over \mathfrak{R} . \mathfrak{S} is a total matrix algebra \mathfrak{M} over a division algebra \mathfrak{D} , which is equivalent to \mathfrak{R} , \mathfrak{C} , or \mathfrak{Q} . Let the canonical basis

$$(4.1) \quad d_h e_{pq}, \quad h = 1, \dots, \delta; p, q = 1, \dots, n,$$

where the d_h are a canonical basis for \mathfrak{D} , and the e_{pq} a canonical basis for \mathfrak{M} , be chosen for \mathfrak{S} . For this choice of basis all the constants of multiplication are rational. Let \mathfrak{S}' denote the algebra with the basis (4.1) over the rational field. We shall prove

LEMMA 2. *A basis for \mathfrak{S}' , $b_1, b_2, \dots, b_\alpha$, $\alpha = \delta n^2$, can be so chosen that the minimum equation of each element b_h is irreducible in the rational field.*

† Cf. L. E. Dickson, *ibid.*, p. 136.

‡ Cf. also L. E. Bush, *Bulletin of the American Mathematical Society*, vol. 38 (1932), pp. 49-51.

If \mathfrak{S}' is a division algebra, that is, if $n = 1$, then the canonical basis noted at the beginning of §2 is a basis of the kind described in the lemma. For, every basis element satisfies one or the other of the equations, $\lambda - 1 = 0$, $\lambda^2 + 1 = 0$, each of which is irreducible in the rational field \mathfrak{F} .

Suppose now that $n > 1$, and suppose that in attempting to choose a basis of the required sort, we have chosen linearly independent elements b_1, b_2, \dots, b_p each of which satisfies an equation irreducible in \mathfrak{F} . Suppose that $p < \alpha$ and that it is impossible to choose another element of \mathfrak{S}' which satisfies an equation irreducible in \mathfrak{F} and which is linearly independent of b_1, b_2, \dots, b_p . Let b_{p+1}, \dots, b_α be chosen in any way to fill out a basis for \mathfrak{S}' . Then the assumption just made implies that every rational linear combination

$$(4.2) \quad \sum_{h=1}^{\alpha} c_h b_h,$$

where at least one of the c_h , $h > p$, is different from zero, satisfies a minimum equation which is reducible in \mathfrak{F} .

Consider the element b_{p+1} of \mathfrak{S}' . It is a matrix of order n with elements in \mathfrak{D} not all of which are zero. Let the r, s position be a position in which a non-zero element of the matrix b_{p+1} appears. This element is of the form $a_0 + a_1 i + a_2 j + a_3 k$, where not all the rational numbers a_h are zero.† Since b_1, \dots, b_α constitute a basis for \mathfrak{S}' , we can, by forming linear combinations (4.2) with $c_{p+1} \neq 0$, produce matrices which have some certain one of the elements $1, i, j$, or k in the r, s position,‡ and which have any arbitrarily chosen rational linear combinations of d_1, \dots, d_s in the remaining positions. Now our assumption implies that every matrix of \mathfrak{S}' which has some certain one of the elements $1, i, j, k$ in the r, s position satisfies a minimum equation which is reducible in \mathfrak{F} . We shall show that this leads to a contradiction.

Consider the so-called companion matrix B , of order n , of the equation $\lambda^n - 2 = 0$,

$$B = \begin{vmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 2 & 0 & 0 & 0 & \cdots & 0 \end{vmatrix}.$$

† It is to be understood that if \mathfrak{D} is \mathfrak{R} , $a_1 = a_2 = a_3 = 0$, and if \mathfrak{D} is \mathbb{C} , $a_2 = a_3 = 0$.

‡ For example, if $a_2 \neq 0$, it is possible to form such matrices with j in the r, s position.

$\lambda^n - 2 = 0$ is the minimum equation of B . Furthermore, if $B' = PBP^{-1}$ is a matrix similar to B , where P is nonsingular with elements in the complex field, then $\lambda^n - 2 = 0$ is also the minimum equation of B' . Now it is fairly evident that a matrix P can be selected so that PBP^{-1} will have a prescribed one of the numbers $1, i, j, k$ in the r, s position. Let u ($1, i, j$, or k) be the element in the r, s position of the matrices constructed in the preceding paragraph. One may verify that, in the several possible cases, the matrices P , listed below, will transform B into the matrix PBP^{-1} , whose element in the r, s position is u .

(1) If $r \neq s - 1$ and $s \neq 1$, $P = I + U_1$, where I is the identity matrix, and U_1 is a matrix with u in the $r, (s-1)$ position and zeros elsewhere.

(2) If $r = s - 1$, and $s \neq 1$, P is a matrix with u in the r, r position, 1's elsewhere on the main diagonal, and zeros in the remaining positions.

(3) If $s = 1$ and $r \neq n$, $P = I + U_3$, where U_3 is a matrix with $u/2$ in the r, n position and zeros elsewhere.

(4) If $s = 1$ and $r = n$, P is a matrix with $u/2$ in the n, n position, 1's elsewhere on the main diagonal, and zeros in the remaining positions.

Now in each of the above cases, P has elements which belong to a field which is isomorphic with the complex field, because $u^2 = 1$ or -1 . Hence, in each of the above cases, the matrix PBP^{-1} satisfies the irreducible (in \mathfrak{F}) equation $\lambda^n - 2 = 0$, and has the number u in the r, s position. This contradicts the previous conclusion that a matrix with u in the r, s position should have a minimum equation reducible in \mathfrak{F} . Consequently, the initial assumption $p < \alpha$ is untenable, and Lemma 2 is proved.

Let it be remarked that if the basis (4.1) is chosen for \mathfrak{S} , then the basis b_1, \dots, b_α of Lemma 2 can be obtained from (4.1) by a rational transformation of basis.

We return now to the consideration of the arbitrary non-nilpotent algebra \mathfrak{A} . Since \mathfrak{A}^* is semisimple it is equivalent to a direct sum of simple algebras $\mathfrak{S}_1, \mathfrak{S}_2, \dots, \mathfrak{S}_\beta$, so that

$$\mathfrak{A} = \mathfrak{A}^* + \mathfrak{Z} = \mathfrak{S}_1 + \mathfrak{S}_2 + \dots + \mathfrak{S}_\beta + \mathfrak{Z}.$$

Each \mathfrak{S}_h has a principal unit e_h , and $e_h e_l = \delta_{hl} e_h$, where δ_{hl} is Kronecker's delta. \mathfrak{Z} can be separated into a sum of $\beta + 1$ linear systems

$$(4.3) \quad e_1 \mathfrak{Z}, e_2 \mathfrak{Z}, \dots, e_\beta \mathfrak{Z}, \mathfrak{Z}',$$

where \mathfrak{Z}' consists of all the elements of \mathfrak{Z} , for which $az = 0$ for every element a of \mathfrak{A}^* .† The linear systems (4.3) are supplementary in their sum, that is, the intersection of any two of them is zero. For, $e_h z_1 = e_l z_2$, $h \neq l$, implies that

† L. E. Dickson, op. cit., pp. 128-130.

$e_h(e_h z_1) = e_h z_1 = 0$; and $e_h z_1 = z'$, where z' is in \mathcal{Z}' , implies that $e_h(e_h z_1) = e_h z_1 = 0$. Consequently, a set of bases for the linear systems (4.3) constitutes a basis for \mathcal{Z} .

Now any system $e_h \mathcal{Z}$ is closed under multiplication on the left by elements of \mathcal{A}^* , in particular, by elements of the simple algebra \mathcal{S}_h , with which it is associated. For, if a^* is any element of \mathcal{A}^* , then $a^*(e_h \mathcal{Z}) = e_h(a^* \mathcal{Z}) \subseteq e_h \mathcal{Z}$. If s_h is in \mathcal{S}_h , and $h \neq l$, then $s_h(e_l \mathcal{Z}) = 0$.

We wish to show that if a rational basis of the type of Lemma 2 is chosen for \mathcal{A}^* , a basis for \mathcal{Z} may be so chosen that the constants of multiplication for the product of a basis element of \mathcal{A}^* by a basis element of \mathcal{Z} , in that order, will be rational numbers. To that end we prove

LEMMA 3. *Let \mathcal{S} be any one of the simple components of \mathcal{A}^* , and let e be the principal unit of \mathcal{S} . Let the canonical basis (4.1) be chosen for \mathcal{S} . Then for this basis of \mathcal{S} , a basis for $e\mathcal{Z}$ can be so chosen that the constants of multiplication for the product of any basis element of \mathcal{S} by any basis element of $e\mathcal{Z}$, in that order, will be rational.*

If there is an element $z_1^{(1)}$ of \mathcal{Z} , for which $e_{11}z_1^{(1)} \neq 0$, choose $e_{11}z_1^{(1)}$ as one of the basis elements of $e\mathcal{Z}$. If there is an element $z_1^{(2)}$ of \mathcal{Z} such that $e_{11}z_1^{(1)}$ and $e_{11}z_1^{(2)}$ are left linearly independent over \mathcal{D} , choose $e_{11}z_1^{(2)}$ as a second basis element. Continue in this manner, choosing as many further elements $e_{11}z_1^{(3)}, \dots, e_{11}z_1^{(r_1)}$ as possible which are such that

$$(4.4) \quad e_{11}z_1^{(1)}, e_{11}z_1^{(2)}, \dots, e_{11}z_1^{(r_1)}$$

are left linearly independent over \mathcal{D} . Then any other element of $e\mathcal{Z}$ of the form $e_{11}z$ is left linearly dependent over \mathcal{D} on (4.4). When the set (4.4) is thus maximal, or if no element $e_{11}z \neq 0$ exists, select an element $z_2^{(1)}$ of \mathcal{Z} which is such that $e_{12}z_2^{(1)}$ is left linearly independent over \mathcal{D} of the elements of (4.4), if such an element $z_2^{(1)}$ exists. Take $e_{12}z_2^{(1)}$ as a basis element of $e\mathcal{Z}$. If there is an element $z_2^{(2)}$ of \mathcal{Z} which is such that $e_{12}z_2^{(1)}$ and $e_{12}z_2^{(2)}$ are left linearly independent over \mathcal{D} of $e_{12}z_2^{(1)}$ and the elements of (4.4), choose $e_{12}z_2^{(2)}$ as one of the basis elements of $e\mathcal{Z}$. When, in the continuation of this process, the set

$$(4.5) \quad e_{12}z_2^{(1)}, e_{12}z_2^{(2)}, \dots, e_{12}z_2^{(r_2)}$$

is as large as possible, or if no such element $e_{12}z_2^{(1)}$ exists, we choose as further basis elements a maximal set

$$e_{13}z_3^{(1)}, e_{13}z_3^{(2)}, \dots, e_{13}z_3^{(r_3)},$$

which, if such elements exist, together with the elements of (4.4) and (4.5)

are left linearly independent over \mathfrak{D} . Continuing in this manner, we finally obtain a set of elements

$$(4.6) \quad e_{1h_l z_{h_l}}^{(m_l)}, \quad l = 1, \dots, \xi; m_l = 1, \dots, \nu_l, \dagger$$

where h_1, h_2, \dots, h_ξ is some subset of $1, 2, \dots, n$.[‡] The elements of (4.6) are left linearly independent over \mathfrak{D} , and moreover, there is no element of $e\mathfrak{B}$ of the form $e_{1h}z$ which is left linearly independent of the elements of (4.6).

Now the elements of the set

$$(4.7) \quad e_{ph_l z_{h_l}}^{(m_l)}, \quad p = 1, 2, \dots, n; l = 1, \dots, \xi; m_l = 1, \dots, \nu_l,$$

are left linearly independent over \mathfrak{D} . For a relation

$$\sum_{\substack{p=1, \dots, n, \\ l=1, \dots, \xi, \\ m_l=1, \dots, \nu_l}} d_{ph_l m_l} e_{ph_l z_{h_l}}^{(m_l)} = 0,$$

where the numbers $d_{ph_l m_l}$ are in \mathfrak{D} , implies, on multiplying on the left by e_{1q} ,

$$(4.8) \quad \sum_{\substack{l=1, \dots, \xi, \\ m_l=1, \dots, \nu_l}} d_{qh_l m_l} e_{1h_l z_{h_l}}^{(m_l)} = 0$$

for every q . But since the elements $e_{1h_l z_{h_l}}^{(m_l)}$ were chosen to be left linearly independent over \mathfrak{D} , (4.8) implies that $d_{qh_l m_l} = 0$ for every q, h_l , and m_l . Thus the elements of (4.7) are left linearly independent over \mathfrak{D} .

Furthermore, the elements of (4.7) constitute a (left) basis for $e\mathfrak{B}$ over \mathfrak{D} , which may be seen as follows. In the first place, every element of $e\mathfrak{B}$ is the product of $e = e_{11} + e_{22} + \dots + e_{nn}$ by an element of \mathfrak{B} , in that order. The existence of an element

$$ez = e_{11}z + e_{22}z + \dots + e_{nn}z$$

of $e\mathfrak{B}$ which is left linearly independent over \mathfrak{D} of the elements (4.7) implies that at least one of the elements $e_{rr}z$ is left linearly independent of (4.7). This implies that $e_{1r}z$ is also left linearly independent over \mathfrak{D} of (4.7); for, if $e_{1r}z$ is a left linear combination of the elements of (4.7), then so is $e_{rr}z$, as may be seen by multiplying $e_{1r}z$ on the left by e_{r1} . But if $e_{1r}z$ is left linearly independent of (4.7), it is left linearly independent of (4.6). This contradicts the hypothesis that (4.6) is a maximal set.

Consequently, the elements

[†] It is assumed that $e\mathfrak{B} \neq 0$; if $e\mathfrak{B} = 0$, Lemma 3 is trivially true.

[‡] If, for instance, there is no element $e_{1h}z$ left linearly independent of (4.4) over \mathfrak{D} , then 2 will not occur among the h_1, \dots, h_ξ .

$$(4.9) \quad d_q e_{phl} z_{hl}^{(ml)}, \\ q = 1, \dots, \delta; p = 1, \dots, n; l = 1, \dots, \zeta; m_l = 1, \dots, \nu_l,$$

constitute a basis for $e\mathfrak{Z}$ over \mathfrak{K} . For this basis of $e\mathfrak{Z}$ it is clear that the constants of multiplication for the product of a canonical basis element of \mathfrak{S} by a basis element of $e\mathfrak{Z}$, in that order, are rational. In fact these constants of multiplication are 0's, 1's, and -1's.

We remark that if \mathfrak{S} is subjected to a rational transformation of basis, from the basis (4.1) to a new basis, and the above basis for $e\mathfrak{Z}$ is left unchanged, then the constants of multiplication for the product of a basis element of \mathfrak{S} by another basis element of \mathfrak{S} , or by a basis element of $e\mathfrak{Z}$, in that order, remain rational. This is true, in particular, for the basis of Lemma 2.

We are now in a position to prove the fundamental theorem on which the extension of the results of §§2 and 3 depends.

THEOREM 4.1. *Let $\mathfrak{A} = \mathfrak{A}^* + \mathfrak{Z}$ be an algebra over \mathfrak{K} , with the radical \mathfrak{Z} , and semisimple component \mathfrak{A}^* . The signatures of $T_1(\mathfrak{A})$, $T_2(\mathfrak{A})$, and ${}_A T(\mathfrak{A}^*)$ are equal.*

Since $\mathfrak{A}^* = \mathfrak{S}_1 + \mathfrak{S}_2 + \dots + \mathfrak{S}_\beta$, and $\mathfrak{A} = \mathfrak{A}^* + \mathfrak{Z}$, we may choose a basis for \mathfrak{A} by choosing bases for $\mathfrak{S}_1, \mathfrak{S}_2, \dots, \mathfrak{S}_\beta$ and \mathfrak{Z} . Let e_1, e_2, \dots, e_β be the respective principal units of $\mathfrak{S}_1, \mathfrak{S}_2, \dots, \mathfrak{S}_\beta$. As previously noted, a basis for \mathfrak{Z} may be chosen to consist of the bases for such of the systems $e_1\mathfrak{Z}, e_2\mathfrak{Z}, \dots, e_\beta\mathfrak{Z}, \mathfrak{Z}'$ as are not zero. Let

$$\begin{aligned} s_1^{(1)}, \dots, s_{\alpha_1}^{(1)}; s_1^{(2)}, \dots, s_{\alpha_2}^{(2)}; \dots; s_1^{(\beta)}, \dots, s_{\alpha_\beta}^{(\beta)}; \\ z_1^{(1)}, \dots, z_{\lambda_1}^{(1)}; z_1^{(2)}, \dots, z_{\lambda_2}^{(2)}; \dots; z_1^{(\beta)}, \dots, z_{\lambda_\beta}^{(\beta)}; z_1, \dots, z_\gamma, \end{aligned}$$

be a basis for \mathfrak{A} , where $s_1^{(h)}, \dots, s_{\alpha_h}^{(h)}$ is a basis for \mathfrak{S}_h , $z_1^{(h)}, \dots, z_{\lambda_h}^{(h)}$ is a basis for $e_h\mathfrak{Z}$, and z_1, \dots, z_γ is a basis for \mathfrak{Z}' , and where it is to be understood that $z_1^{(h)}, \dots, z_{\lambda_h}^{(h)}$ are absent if $e_h\mathfrak{Z} = 0$, and similarly for the z_p , if $\mathfrak{Z}' = 0$.

Consider any one of the simple algebras \mathfrak{S}_h . By Lemma 2 the basis elements $s_1^{(h)}, \dots, s_{\alpha_h}^{(h)}$ can be taken to be such that each $s_m^{(h)}$ satisfies an equation which is irreducible in the rational field \mathfrak{F} . Now $s_m^{(h)}z = 0$ for every element z which is in $e_1\mathfrak{Z} + \dots + e_{h-1}\mathfrak{Z} + e_{h+1}\mathfrak{Z} + \dots + e_\beta\mathfrak{Z} + \mathfrak{Z}'$. By Lemma 3, if $e_h\mathfrak{Z} \neq 0$, the basis $z_1^{(h)}, \dots, z_{\lambda_h}^{(h)}$ can be so chosen that the constants of multiplication for a product $s_m^{(h)}z_i^{(h)}$ are rational.

Consider now the first matrix ${}_A R(s_q^{(h)})$ of any one of the basis elements of \mathfrak{S}_h , where ${}_A R(s_q^{(h)})$ denotes the first matrix of $s_q^{(h)}$ in the first matrix representation of \mathfrak{A} :

$$(4.10) \quad \mathfrak{A}R(s_q^{(h)}) = \left\| \begin{array}{cccccc} 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \mathfrak{E}_h R(s_q^{(h)}) & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & \cdots & e_{h\beta} R(s_q^{(h)}) & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{array} \right\|,$$

where the 0's stand for blocks of zeros, $\mathfrak{E}_h R(s_q^{(h)})$ (occurring in the h th block down and the h th block over) is the first matrix of $s_q^{(h)}$ in the first matrix representation of \mathfrak{E}_h , and where $e_{h\beta} R(s_q^{(h)})$ is a matrix of order λ_h , whose elements are the constants of multiplication of products of $s_q^{(h)}$ by basis elements of $e_{h\beta}$. Since \mathfrak{E}_h has a principal unit, the matrices $\mathfrak{E}_h R(s_q^{(h)})$, ($q = 1, 2, \dots, \alpha_h$), are linearly independent.† Hence the matrices $\mathfrak{A}R(s_q^{(h)})$ are linearly independent, and therefore there is a simple ring isomorphism between the $\mathfrak{A}R(s_q^{(h)})$ and the $s_q^{(h)}$.†

The matrix $\mathfrak{A}R(e_h)$ is the matrix (4.10), where $\mathfrak{E}_h R(s_q^{(h)})$ and $e_{h\beta} R(s_q^{(h)})$ are identity matrices of orders α_h and λ_h , respectively, since e_h is a left-hand principal unit for \mathfrak{E}_h and $e_{h\beta}$. Now each $s_q^{(h)}$, ($q = 1, \dots, \alpha_h$), satisfies an equation

$$f_q(x) = x^n + c_{n-1,q}x^{n-1} + \cdots + c_{1,q}x + c_{0,q} = 0,$$

irreducible in \mathfrak{F} , when $c_{0,q}$ is replaced by $c_{0,q}e_h$. Therefore, $\mathfrak{A}R(s_q^{(h)})$ also satisfies $f_q(x) = 0$, if $c_{0,q}$ is replaced by $c_{0,q} \cdot \mathfrak{A}R(e_h)$. Hence, $\mathfrak{E}_h R(s_q^{(h)})$ and $e_{h\beta} R(s_q^{(h)})$ also satisfy $f_q(x) = 0$, when $c_{0,q}$ is replaced respectively by I_{α_h} and I_{λ_h} , the identity matrices of orders α_h and λ_h . Since the elements of each of $\mathfrak{E}_h R(s_q^{(h)})$ and $e_{h\beta} R(s_q^{(h)})$ are rational, and since $f_q(x)$ is irreducible in \mathfrak{F} , $f_q(x)$ is the minimum function of each of $\mathfrak{E}_h R(s_q^{(h)})$ and $e_{h\beta} R(s_q^{(h)})$, because the minimum function of a matrix divides any polynomial which vanishes for that matrix. The characteristic function of $\mathfrak{E}_h R(s_q^{(h)})$ is therefore a power $[f_q(x)]^\phi$ of its minimum function $f_q(x)$, and hence the trace of $\mathfrak{E}_h R(s_q^{(h)})$ is equal to $\phi c_{n-1,q}$. Likewise the characteristic function of $e_{h\beta} R(s_q^{(h)})$ is a power $[f_q(x)]^\psi$ of $f_q(x)$, and the trace of $e_{h\beta} R(s_q^{(h)})$ is $\psi c_{n-1,q}$. Now the first trace of $s_q^{(h)}$ relative to \mathfrak{A} , $\mathfrak{A}t_1(s_q^{(h)})$, is the trace of $\mathfrak{A}R(s_q^{(h)})$, which is equal to the sum of the traces of $\mathfrak{E}_h R(s_q^{(h)})$ and $e_{h\beta} R(s_q^{(h)})$. Hence

† Cf. L. E. Dickson, op. cit., p. 34.

$$\begin{aligned}
 {}_{\mathfrak{A}}t_1(s_q^{(h)}) &= t({}_{\mathfrak{A}}R(s_q^{(h)})) = (\phi + \psi)c_{n-1,q} \\
 (4.11) \qquad &= \left(1 + \frac{\psi}{\phi}\right)\phi c_{n-1,q} = \left(1 + \frac{\lambda_h}{\alpha_h}\right) \cdot t({}_{\mathfrak{A}}R(s_q^{(h)})).
 \end{aligned}$$

We have arrived at (4.11) on the assumption that $e_h\beta \neq 0$. However, if $e_h\beta = 0$, then ${}_{e_h}R(s_q^{(h)})$ does not appear in the matrix (4.10). Further, $\lambda_h = 0$. From this it is apparent that (4.11) also holds if $e_h\beta = 0$.

Now $t({}_{\mathfrak{A}}R(s_q^{(h)})) = {}_{\mathfrak{A}}t_1(s_q^{(h)})$, the first trace of $s_q^{(h)}$ relative to \mathfrak{A}^* . It is known that $t_1(a) = t_2(a)$, for every element a of a semisimple algebra (cf. R1). Hence we may write (4.11) as

$$(4.12) \qquad {}_{\mathfrak{A}}t_1(s_q^{(h)}) = \theta_h [{}_{\mathfrak{A}}t(s_q^{(h)})],$$

where $\theta_h = 1 + \lambda_h/\alpha_h > 0$. Note that θ_h is the same for every element $s_q^{(h)}$ of \mathfrak{E}_h .

Since every element of \mathfrak{E}_h is a linear combination of $s_1^{(h)}, s_2^{(h)}, \dots, s_{\alpha_h}^{(h)}$ with coefficients in \mathfrak{A} , and since the trace of a linear combination of elements is equal to the same linear combination of the traces of those elements, it follows from (4.12) that

$$(4.13) \qquad {}_{\mathfrak{A}}T_1(\mathfrak{E}_h) = \theta_h [{}_{\mathfrak{E}_h}T(\mathfrak{E}_h)].$$

Since \mathfrak{E}_h was chosen arbitrarily, (4.13) holds for every h . Of course θ_h may change with h . Hence we may write

$$(4.14) \quad T_1(\mathfrak{A}) = \left\| \begin{array}{cccccc} \theta_1 \cdot {}_{\mathfrak{E}_1}T(\mathfrak{E}_1) & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \theta_2 \cdot {}_{\mathfrak{E}_2}T(\mathfrak{E}_2) & \cdots & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \theta_{\beta} \cdot {}_{\mathfrak{E}_{\beta}}T(\mathfrak{E}_{\beta}) & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{array} \right\|.$$

The signature, $\sigma(T_1(\mathfrak{A}))$, of $T_1(\mathfrak{A})$ is the sum of the signatures of the matrices $\theta_h [{}_{\mathfrak{E}_h}T(\mathfrak{E}_h)]$. Since $\theta_h > 0$, the signature of $\theta_h [{}_{\mathfrak{E}_h}T(\mathfrak{E}_h)]$ is the same as the signature of ${}_{\mathfrak{E}_h}T(\mathfrak{E}_h)$. But, since \mathfrak{A}^* is the direct sum of the \mathfrak{E}_h , the sum of the signatures of the ${}_{\mathfrak{E}_h}T(\mathfrak{E}_h)$ is exactly the signature of ${}_{\mathfrak{A}}T(\mathfrak{A}^*)$. This proves Theorem 4.1 for the signatures of $T_1(\mathfrak{A})$ and ${}_{\mathfrak{A}}T(\mathfrak{A}^*)$ for a particular basis of \mathfrak{A} .

Under a transformation of basis of \mathfrak{A} of matrix C , $T_1(\mathfrak{A})$ is transformed into $CT_1(\mathfrak{A})C^T$, and $\sigma(T_1(\mathfrak{A}))$ is invariant. Now the semisimple algebra \mathfrak{A}^* of \mathfrak{A} is not unique, but any two such components are equivalent. Since the discriminant matrices depend only on the constants of multiplication, the

fact that \mathfrak{A}^* and \mathfrak{A}^{**} are equivalent is sufficient to insure the equality of their discriminant matrices for isomorphic bases. Since, further, a transformation of basis of \mathfrak{A}^* does not change the signature of ${}_q T(\mathfrak{A}^*)$, it follows that if $\sigma(T_1(\mathfrak{A})) = \sigma(T(\mathfrak{A}^*))$ for one choice of basis of \mathfrak{A} , the like is true for all bases.

Now it should be evident that if we should make a right-hand decomposition of Z into the $\beta+1$ linear systems

$$\mathfrak{Z}e_1, \mathfrak{Z}e_2, \dots, \mathfrak{Z}e_\beta, \mathfrak{Z}'',$$

the analogue of Lemma 3 can be stated and proved in a "right-hand" way. Then the above proof can be carried out in a precisely analogous manner for the equality of the signatures of $T_2(\mathfrak{A})$ and ${}_q T(\mathfrak{A}^*)$, by use of the second matrices ${}_q S(s_q^{(h)})$ of the elements of \mathfrak{S}_h . The constants θ'_h will of course not be necessarily the same as the corresponding θ_h . This completes the proof of Theorem 4.1.

We pause briefly to note two corollaries to the proof of Theorem 4.1, which have no direct bearing on the problem of this paper, but which are of interest in their own right. It has been shown that the first and second discriminant matrices of an algebra, relative to a given basis, are not, in general, equal (cf. R1). However, from Theorem A of §1 they have the same rank, and by Theorem 4.1 they have the same signature. Moreover,

COROLLARY 4.11. *Let \mathfrak{P} be a primary algebra over \mathfrak{K} , with the radical \mathfrak{Z} . Then $T_1(\mathfrak{P})$ is equal to a scalar times $T_2(\mathfrak{P})$.*

Since \mathfrak{P} is primary, it is equivalent to the sum of its radical \mathfrak{Z} and a simple algebra \mathfrak{S} . For the particular basis B used in the proof of Theorem 4.1, we have from (4.14)

$$(4.15) \quad T_1(\mathfrak{P}) = \theta \cdot [{}_S T(\mathfrak{S}) + O],$$

where θ is a nonzero scalar matrix and O is a zero matrix of order p , the order of \mathfrak{Z} . Likewise, for another similarly chosen basis B' for \mathfrak{P} , we would have

$$(4.16) \quad T_2(\mathfrak{P}) = \theta' \cdot [{}_S T(\mathfrak{S}) + O],$$

where θ' is a nonzero scalar matrix. Now the bases B and B' for which (4.14) and (4.15) hold, differ only in the choice of basis for the radical \mathfrak{Z} . It is therefore possible to make a transformation of basis of \mathfrak{P} from B to B' , by a transformation whose matrix Q is of the form

$$(4.17) \quad Q = I_\alpha + M,$$

where I_α is the identity matrix of order α , α being the order of \mathfrak{S} . Hence, relative to the basis B' , $T_1(\mathfrak{P})$ becomes

$$(4.18) \quad T'_1(\mathfrak{P}) = Q T_1(\mathfrak{P}) Q^T = Q \theta [\varepsilon T(\mathfrak{E}) + O] Q^T = \theta \cdot [\varepsilon T(\mathfrak{E}) + O].$$

Hence, relative to the basis B' , $T'_1(\mathfrak{P}) = (\theta/\theta') T'_2(\mathfrak{P})$. Since $T_1(\mathfrak{P})$ and $T_2(\mathfrak{P})$ are transformed cogrediently under transformations of basis of \mathfrak{P} , it follows that $T_1(\mathfrak{P}) = (\theta/\theta') T_2(\mathfrak{P})$ for every basis of \mathfrak{P} .

COROLLARY 4.12. *Let $\mathfrak{A} = \mathfrak{A}^* + \mathfrak{J}$ be an algebra over \mathfrak{K} , with radical \mathfrak{J} and semisimple component \mathfrak{A}^* . It is possible to choose a basis for \mathfrak{A} such that $T_1(\mathfrak{A})$, $T_2(\mathfrak{A})$, and ${}_A T(\mathfrak{A}^*)$ simultaneously assume a diagonal form.*

A basis for a semisimple algebra for which the discriminant matrix assumes a diagonal form has been called by MacDuffee (M2) a normal basis. Let $\mathfrak{A} = \mathfrak{E}_1 + \mathfrak{E}_2 + \cdots + \mathfrak{E}_s + \mathfrak{J}$, where the \mathfrak{E}_h are simple algebras. Let a basis for \mathfrak{A} be chosen to consist of normal bases for the \mathfrak{E}_h and any basis for \mathfrak{J} . From the proof of Theorem 4.1, or from Corollary 4.11 we have

$${}_A T_1(\mathfrak{E}_h) = \theta_h [\varepsilon_h T(\mathfrak{E}_h)], \quad {}_A T_2(\mathfrak{E}_h) = \theta'_h [\varepsilon_h T(\mathfrak{E}_h)], \quad h = 1, \dots, s,$$

so that for the present basis choice, ${}_A T_1(\mathfrak{E}_h)$ and ${}_A T_2(\mathfrak{E}_h)$ are diagonal matrices because $\varepsilon_h T(\mathfrak{E}_h)$ is such. But $T_1(\mathfrak{A})$ is a direct sum of the ${}_A T_1(\mathfrak{E}_h)$ and a zero matrix of order equal to the order of \mathfrak{J} . Hence $T_1(\mathfrak{A})$ is a diagonal matrix. Similarly, $T_2(\mathfrak{A})$, relative to this basis, is a diagonal matrix.

5. Extension to an arbitrary algebra. With the aid of Theorem 4.1 the extension of Theorems 3.1 and 3.2 is readily made. Let $\mathfrak{A} = \mathfrak{A}^* + \mathfrak{J}$ be an arbitrary algebra over the real field. First, let \mathfrak{A} be non-nilpotent, that is, assume $\mathfrak{A}^* \neq 0$. A nilpotent subalgebra \mathfrak{L} of \mathfrak{A} of maximal order will clearly be the sum of the radical \mathfrak{J} and a nilpotent subalgebra \mathfrak{L}^* of \mathfrak{A}^* of maximal order. A complete set of primitive idempotents of \mathfrak{A}^* also constitutes a complete set of \mathfrak{A} . Let μ denote the number of nonnegative terms in a diagonal form of $T_1(\mathfrak{A})$ [or $T_2(\mathfrak{A})$]. That is, $\mu = n + (\sigma - \rho)/2$, where n is the order, σ the signature, ρ the rank of $T_1(\mathfrak{A})$ [or $T_2(\mathfrak{A})$]. Also $\mu = \mu^* + \omega$, where μ^* is the index of inertia of ${}_A T_1(\mathfrak{A}^*)$ [or ${}_A T_2(\mathfrak{A}^*)$], and where ω is the common nullity of $T_1(\mathfrak{A})$ and $T_2(\mathfrak{A})$. From Theorem 4.1 it follows that the indices of inertia of ${}_A T_1(\mathfrak{A}^*)$, ${}_A T_2(\mathfrak{A}^*)$, and ${}_A T(\mathfrak{A}^*)$ are equal. From this and from Theorem 3.1 $\mu^* = \epsilon + \chi^*$, where ϵ is the number in a complete set of primitive idempotents of \mathfrak{A}^* [or \mathfrak{A}], and χ^* is the order of \mathfrak{L}^* . Therefore, $\mu = \epsilon + \chi$ where χ is the order of \mathfrak{L} .

If \mathfrak{A} is nilpotent, then $\epsilon = 0$, $T_1(\mathfrak{A}) = T_2(\mathfrak{A}) = 0$, $\mu = n$, and the relation $\mu = \epsilon + \chi$ is trivially true.

This completes the proof of the general

THEOREM 5.1. *Let \mathfrak{A} be an arbitrary algebra of order n over \mathfrak{K} . Let μ ($\mu = n + (\sigma - \rho)/2$, where ρ and σ are, respectively, the rank and signature of $T_1(\mathfrak{A})$ [$T_2(\mathfrak{A})$]), be the number of nonnegative terms in the diagonal of a diagonal*

form of $T_1(\mathfrak{A})$ [$T_2(\mathfrak{A})$]. Then μ is equal to the order of a nilpotent subalgebra of \mathfrak{A} of maximal order, plus the number in a complete set of primitive idempotents of \mathfrak{A} .

Also, from Theorem 3.2 follows

THEOREM 5.2. *Let \mathfrak{A} and μ have the same significance as in Theorem 5.1. Then μ is equal to the order of a subalgebra of \mathfrak{A} of minimum order which contains a complete set of primitive idempotents of \mathfrak{A} , and which has, as its radical, a nilpotent subalgebra of \mathfrak{A} of maximal order.*

It may be remarked that a nilpotent subalgebra of \mathfrak{A} of maximal order is also obviously maximal in the sense of the calculus of complexes.

6. Specialization to the Borchardt-Jacobi Theorem. To demonstrate how Theorem 5.1 (or 5.2) is a generalization of part II of the B. J. Theorem, let us specialize \mathfrak{A} to a polynomial algebra. Let $p(x)=0$ be a polynomial equation of degree n with real coefficients and with leading coefficient unity. Let \mathfrak{X} be the polynomial algebra over \mathfrak{R} generated by $p(x)$. Over \mathfrak{R} , $p(x)$ can be decomposed into powers of distinct irreducible factors thus:

$$p(x) = \prod (x - a_i)^{h_i} (x^2 + b_j x + c_j)^{k_j}, \quad i = 1, \dots, r; j = 1, \dots, s,$$

where the a_i , b_j , and c_j are in \mathfrak{R} . It is known that \mathfrak{X} is equivalent to a direct sum of the $r+s$ polynomial algebras generated by the $(x-a_i)^{h_i}$ and the $(x^2+b_j x+c_j)^{k_j}$ (cf. R2). Since \mathfrak{X} is commutative and has a principal unit, the number of primitive idempotents of \mathfrak{X} is equal to the number of primary component algebras in the direct sum decomposition of \mathfrak{X} .† Hence $\epsilon = r+s$. Again, because \mathfrak{X} is commutative, the nilpotent subalgebra of \mathfrak{X} of maximal order coincides with the radical of \mathfrak{X} , both consisting of all the nilpotent elements of \mathfrak{X} . Now the order of the radical of \mathfrak{X} is

$$\chi = \sum_{i=1}^r (h_i - 1) + 2 \sum_{j=1}^s (k_j - 1) = n - (r + 2s).$$

Hence the rank ρ of the discriminant matrix of \mathfrak{X} is $r+2s$.

By Theorem 5.1 the signature of $T(\mathfrak{X})$ must be equal to $2(\mu-n)+\rho = 2(\chi+\epsilon-n)+\rho$. But $2(\chi+\epsilon-n)+\rho = 2[r+s+n-(r+2s)-n]+r+2s=r$. Hence the signature of $T(\mathfrak{X})$ is equal to r , the number of distinct real roots of $p(x)=0$. This result is part II of the B. J. Theorem.

† See, for instance, G. Scorza, *Sulle algebre riducibili*, Rendiconti del Seminario Matematico delle Università di Roma, (4), vol. 1 (1937), pp. 188-189.

GEOMETRIC ASPECTS OF RELATIVISTIC DYNAMICS*

BY

L. A. MAC COLL

INTRODUCTION

1. Kasner has studied the three-parameter families of trajectories of a particle moving in a plane under forces which are functions of position only, and has shown that all such families of curves, each particular family corresponding to a particular field of force, possess certain common geometrical properties which distinguish them from three-parameter families of curves defined in other ways.† He and his students have also studied a variety of other problems concerning families of trajectories of particles, but in all of this work it has been assumed that the particles obey the laws of Newtonian dynamics. So far there do not seem to have been any parallel investigations concerning the trajectories of particles obeying the laws of special relativistic dynamics.

For the sake of brevity, we shall call a particle obeying the laws of Newtonian dynamics a classical particle, and we shall call a particle obeying the laws of special relativistic dynamics a relativistic particle.

This article deals primarily with the problem of determining a set of geometrical properties which is characteristic of the families of trajectories of a relativistic particle moving in a plane under forces which are functions of position only. Whereas Kasner found that in the classical case the families of trajectories are characterized by a certain set of five properties, we find that in the relativistic case there are six characteristic properties.‡ Four of these correspond to four of the properties given by Kasner for the classical case, and resemble the latter in various degrees, while the remaining two properties have no classical analogues.

In the concluding sections of the article we deal with some other problems concerning trajectories of relativistic particles, most of the considerations being confined to the case of motion in a plane. In particular, we study the determination of the field of force by the properties of the family of trajectories,

* Presented to the Society, October 29, 1938; received by the editors January 11, 1939.

† These Transactions, vol. 7 (1906), pp. 401-424; also *Differential-Geometric Aspects of Dynamics*, American Mathematical Society Colloquium Publications, vol. 3, New York, 1913, pp. 9-17.

‡ When this paper was presented to the Society, on October 29, 1938 (Bulletin of the American Mathematical Society, abstract 44-9-397), it was announced that the families of trajectories can be characterized by a set of seven properties. It has since been found that one of those properties is a consequence of the others.

we investigate point transformations which transform families of trajectories into families of trajectories, and we consider the properties of certain special families of trajectories which are called natural families. (A natural family of trajectories is the family of possible trajectories of a particle moving in a conservative field of force with a prescribed value of the total energy.)

In many places the detailed proofs of the results will be omitted; for these proofs depend, for the most part, upon entirely elementary and straightforward, but tedious, calculations.

THE DIFFERENTIAL EQUATION DEFINING THE FAMILY OF TRAJECTORIES

2. We consider a relativistic particle, having rest-mass m_0 , moving in a plane under a force which is a function of position only. If x and y are the rectangular coordinates of the particle with respect to a fixed set of axes, and if $X(x, y)$ and $Y(x, y)$ are, respectively, the x -component and the y -component of the force, the differential equations of motion of the particle can be written in the form

$$(1) \quad \begin{aligned} \frac{d}{dt} \left[\dot{x} \left(1 - \frac{\dot{x}^2 + \dot{y}^2}{c^2} \right)^{-1/2} \right] &= \frac{1}{m_0} X(x, y) \equiv \phi(x, y), \\ \frac{d}{dt} \left[\dot{y} \left(1 - \frac{\dot{x}^2 + \dot{y}^2}{c^2} \right)^{-1/2} \right] &= \frac{1}{m_0} Y(x, y) \equiv \psi(x, y). \end{aligned}$$

Here, of course, c denotes the speed of light, and the dots indicate total differentiation with respect to the time t . If both ϕ and ψ are identically zero, the family of trajectories is merely the two-parameter family of straight lines in the plane. We explicitly exclude this degenerate case from all of our considerations. We shall assume that the functions ϕ and ψ are of class C^2 , if not throughout the entire plane, at least throughout a certain open region to which our considerations are restricted.*

We first obtain the differential equation defining the family of possible trajectories, by eliminating the time from equations (1) in the usual way. The result is the equation

$$(2) \quad y''' = -F + Gy'' + Hy'^2 + F(1 + Ky'^2)^{1/2},$$

where

$$(3) \quad \begin{aligned} F &= \frac{1}{2c^4} (1 + y'^2)(\psi - \phi y')(\phi + \psi y'), & H &= -\frac{3\phi}{\psi - \phi y'}, \\ G &= \frac{\psi_x + (\psi_y - \phi_x)y' - \phi_y y'^2}{\psi - \phi y'}, & K &= \frac{4c^4}{(1 + y'^2)(\psi - \phi y')^2}. \end{aligned}$$

* Many of our results are valid under conditions which are slightly broader than these. The minimum conditions under which the conclusions hold cannot be stated in any simple form.

The primes indicate total differentiation with respect to x ; and $\phi_x = \partial\phi/\partial x$, and so on. The positive value of the square root in the last term of (2) is the significant one; and wherever square roots appear in the following work it is to be understood, unless the contrary is explicitly indicated, that the positive values are intended. We note the identity

$$(4) \quad FK + 2H/3 = \frac{2y'}{1 + y'^2}.$$

As may be seen by letting c tend to infinity, the equation which corresponds to (2) in the classical case is $y''' = Gy'' + Hy'^2$, G and H being given by the above formulas. We see that, for a given field of force, the family of trajectories is independent of the rest-mass of the particle in the classical case, but not in the relativistic case.

Equation (2) is not an arbitrary differential equation of the third order.* On the contrary, the equation is entirely special in respect to the way in which the derivatives are involved, and it is somewhat special in respect to the way in which x and y are involved. Hence, regardless of the forms of the functions ϕ and ψ , the family of curves defined by (2) must possess certain special geometrical properties, corresponding to the special features of the form of the equation. Our immediate problem is to discover these characteristic properties.

THE CHARACTERISTIC PROPERTIES OF THE FAMILY OF TRAJECTORIES

3. Following Kasner's procedure, we begin by considering the trajectories which pass through a fixed point O : (x, y) in the direction determined by a fixed value of y' , the lineal element (x, y, y') being such that, for it, F , G , and H are all finite, and F and H are not zero.† These curves form a one-parameter family, the different curves having different curvatures at the point O . Considering each of the curves of this family, we construct the parabola which osculates the curve at the point O . Finally, we consider the locus Γ_1 of the foci of these parabolas.

For convenience in discussing the curve Γ_1 and certain other curves, we introduce two auxiliary systems of rectangular coordinates with their origins at the point O . The one, (ξ, η) , system is such that the ξ -axis and η -axis are

* By an arbitrary differential equation of the third order we mean an equation of the form $y''' = f(x, y, y', y'')$, where the right-hand member is an arbitrary function of the four arguments indicated.

† In order to satisfy the condition $H \neq 0$, it may be necessary to make an adjustment of the coordinate system. We may as well assume that the adjustment of the coordinate system is such that ψ also does not vanish at the point O .

parallel to the x -axis and y -axis, respectively. The other, (u, v) , system is such that the u -axis is the common tangent, at O , of the ∞^1 trajectories we are considering. The orientations of both of these sets of axes are the same as that of the (x, y) set. The relation between the auxiliary coordinate systems is represented by the equations

$$\xi + y'\eta = (1 + y'^2)^{1/2}u, \quad -y'\xi + \eta = (1 + y'^2)^{1/2}v.$$

The focus of the parabola determined by the differential element of the third order (x, y, y', y'', y''') has the coordinates

$$\xi = \frac{3}{2} y'' \frac{(1 + y'^2)y''' - 6y'y''^2}{(1 + y'^2)y'''^2 - 6y'y''^2y''' + 9y''^4},$$

$$\eta = \frac{3}{2} y'' \frac{(1 + y'^2)y'y''' + 3(1 - y'^2)y''^2}{(1 + y'^2)y'''^2 - 6y'y''^2y''' + 9y''^4}.$$

The equation of the curve Γ_1 is obtained by eliminating y'' and y''' from these equations and equation (2). We find that the resulting equation, written in terms of the coordinates (u, v) , is

$$(5) \quad (u^2 + v^2)^2 [G(u^2 + v^2) - 2u_0u - 4v_0v/3] - \frac{1}{4F} (1 + y'^2)^{3/2}v [G(u^2 + v^2) - 2u_0u - 2v_0v]^2 = 0,$$

where

$$(6) \quad u_0 = (3/4)(1 + y'^2)^{1/2}, \quad v_0 = (1/4)(1 + y'^2)^{1/2}[3y' - (1 + y'^2)H].$$

The curve Γ_1 is a quintic or a sextic according as G is, or is not, zero.

Let a be an arbitrarily chosen positive constant. The inverse of the curve Γ_1 with respect to the circle $u^2 + v^2 = a^2$ is the cubic Γ'_1 represented by the equation

$$a^4[Ga^2 - 2u_0u - 4v_0v/3] - \frac{1}{4F} (1 + y'^2)^{3/2}v[Ga^2 - 2u_0u - 2v_0v]^2 = 0.$$

This cubic can be obtained from the particular cubic Γ_0 represented by the equation

$$(7) \quad a^2(u + 2v/3) + v(u + v)^2 = 0$$

by means of the affine transformation

$$(8) \quad u \rightarrow \frac{A}{a} \left(u - \frac{Ga^2}{2u_0} \right), \quad v \rightarrow \frac{Av_0}{au_0} v,$$

where

$$(9) \quad A = \frac{3^{1/2}c^2(1 + y'^2)^{1/2}}{2(\phi + \psi y')}$$

We observe that the cubic Γ_1' passes through the point O when, and only when, G is zero. This is the case in which Γ_1 reduces to a quintic. If, and only if, the field of force is given by the equations

$$\phi = a_1 + a_2x, \quad \psi = a_3 + a_2y,$$

where the a 's are constants, the cubic Γ_1' always passes through the corresponding point O . Various physically important fields of force satisfy this condition.

The cubic Γ_1' has the three asymptotes represented by the equations

$$v = 0, \quad u + \frac{v_0}{u_0}v = \frac{Ga^2}{2u_0} \pm 3^{-1/2} \frac{a^2}{A}.$$

The curve has three real branches, one, and only one, of which is asymptotic to both of the parallel asymptotes and is not asymptotic to the third asymptote, $v=0$. Let us call this particular branch the transverse branch of Γ_1' .

Because the square root in the last term of equation (2) is positive, and because the significant values of y'' are all of one sign, as y'' varies the focus of the osculating parabola does not describe the entire curve represented by equation (5), but only a certain arc of the curve. When y'' approaches zero, the coordinates of the inverse of the focus of the osculating parabola approach the values

$$u = \frac{Ga^2}{2u_0}, \quad v = 0.$$

It follows from this fact and some simple continuity considerations that the foci of the osculating parabolas lie on an arc of Γ_1 which is the inverse of a part of the transverse branch of Γ_1' .

Hence we can state the first property of the family of trajectories in the following form:

PROPERTY I. (1) *If, for each of the ∞^1 trajectories passing through a given point in a given direction, we construct the parabola which osculates the trajectory at the given point, the locus Γ_1 of the foci of these parabolas is the inverse of a cubic Γ_1' with respect to the circle $u^2 + v^2 = a^2$, where a is an arbitrary positive constant. The cubic Γ_1' can be obtained from the particular cubic Γ_0 represented by equation (7), by means of an affine transformation of the form (8), where u_0 is given by the first of equations (6), and where A , G , and v_0 are functions of x , y ,*

and y' , and are independent of a . (2) More particularly, the foci of the osculating parabolas lie on an arc of Γ_1 which is the inverse of a part of the transverse branch of Γ'_1 .

The calculations which establish Property I can be reversed unambiguously, and in this way we get a converse of the property. We find that if a three-parameter family of plane curves possesses Property I, the defining differential equation is of the form (2), where now F , G , and H are some functions of x , y , and y' , K is defined by equation (4), and the square root in the final term has its positive value. If a family of curves has the first part of Property I, but not necessarily the second part, the defining differential equation is as just described, except that the sign of the square root is not determined.

We see that Property I is characteristic of families of curves defined by differential equations which have the structure of equation (2) as regards y'' and y''' , the functions (of x , y , and y') F , H , and K being subject to the restriction (4).

It is of interest to consider the relations between these results and the corresponding results for the classical case given by Kasner.* The curve which corresponds in the classical case to our curve Γ_1 is a circle, or a straight line, according as G is not, or is, zero. The inverse of this curve with respect to the circle $u^2 + v^2 = a^2$ is the straight line represented by the equation

$$2u_0u + 2v_0v = Ga^2.$$

This line is parallel to the parallel asymptotes of Γ'_1 , and is midway between them.

It will be observed that if we let c tend to infinity, the curve Γ_1 degenerates, not into the classical circle, but into that circle taken twice, together with the line $v=0$. We can see without difficulty that the second circle and the line $v=0$ constitute the degenerate form of the nonsignificant part of Γ_1 (the part of Γ_1 formed by the inverses of the points of Γ'_1 which do not lie on the transverse branch).

4. The tangent at the point O to the line of force passing through that point is represented by the equation

$$v = \frac{\psi - \phi y'}{\phi + \psi y'} u.$$

As we have seen, the slope of the parallel asymptotes of Γ'_1 is $-u_0/v_0$ in the

* It is understood, of course, that here and elsewhere we are comparing the properties of the family of relativistic trajectories with the properties of the family of classical trajectories in the same field of force.

uv -coordinate system. It readily follows from equations (3) and (6) that

$$-\frac{u_0}{v_0} = -\frac{\psi - \phi y'}{\phi + \psi y'}.$$

Hence we have a second property of the family of trajectories, which can be stated as follows:

PROPERTY II. *The cubic Γ'_1 which corresponds, according to Property I, to a lineal element (x, y, y') is such that the lineal element bisects the angle between the direction of the parallel asymptotes of Γ'_1 and a certain direction which is fixed for the given point O (the direction of the force acting at O).*

Conversely, it is easily shown that if a family of curves possessing Property I also possesses Property II, the function $H(x, y, y')$ in the defining differential equation must be of the form

$$(10) \quad H = \frac{3}{y' - \omega(x, y)},$$

where $\omega(x, y)$ is the slope of the direction, associated with the point (x, y) , which is referred to in the statement of the property.

Property II is very closely related to the second property in Kasner's set. The remarks previously made concerning the relation between Property I and Kasner's corresponding property will suffice to make this connection clear.

5. The point P on the u -axis midway between the parallel asymptotes of the curve Γ'_1 has the coordinates $u = Ga^2/2u_0$, $v = 0$. If, at the point O , we have the relations $\psi_x = \psi_y - \phi_x = \phi_y = 0$, the point P coincides with O for all values of y' ; otherwise, as y' varies the point P describes a certain curve Γ_2 . We readily find that Γ_2 is represented by the equation

$$(\xi^2 + \eta^2)(\psi\xi - \phi\eta) = (2a^2/3) [\psi_x\xi^2 + (\psi_y - \phi_x)\xi\eta - \phi_y\eta^2].$$

The inverse of Γ_2 with respect to the circle $\xi^2 + \eta^2 = a^2$, a being the constant used in defining Γ'_1 , is represented by the equation

$$\psi\xi - \phi\eta = (2/3) [\psi_x\xi^2 + (\psi_y - \phi_x)\xi\eta - \phi_y\eta^2].$$

Thus we have

PROPERTY III. (1) *Either the point P on the u -axis midway between the parallel asymptotes of Γ'_1 coincides with O for all values of y' , or, as y' varies, P describes a curve Γ_2 , which is the inverse, with respect to the circle $\xi^2 + \eta^2 = a^2$, of a conic Γ'_2 passing through the point O . (2) If the conic Γ'_2 exists,* its tangent at O has the direction, fixed for O , referred to in the statement of Property II.*

* We consider that the conic does not exist if P coincides with O for all values of y' .

Conversely, if a family of curves possessing Property I also possesses the first part of Property III, the function $G(x, y, y')$ in the defining differential equation has the form

$$(11) \quad G = \frac{\mu_1 + \mu_2 y' + \mu_3 y'^2}{\omega_1 - y'},$$

where μ_1, μ_2, μ_3 , and ω_1 are functions of x and y . If a family of curves possessing Properties I and II also possesses both parts of Property III, we have the relation $\omega_1 = \omega$, where ω is the function introduced in connection with the converse of Property II.

If we have the relations $\phi = \Phi_x, \psi = \Phi_y$, where Φ is some function of x and y , we say that the field of force is conservative. We note that if, and only if, the field of force is conservative, the conic Γ'_2 , when it exists, is always either a rectangular hyperbola or a pair of perpendicular straight lines. The only conservative fields of force for which the conic never exists are those derived from functions Φ of the form

$$\Phi = a_1 + a_2 x + a_3 y + a_4 (x^2 + y^2),$$

where the a 's are constants.

If ϕ and ψ are, respectively, the real and the imaginary parts of an analytic function of $x+iy$, we have what Lecornu has called an analytic field of force. We see that if, and only if, the field of force is analytic, the conic Γ'_2 when it exists is always a circle. The only analytic fields of force for which the conic never exists are those for which the expressions $\phi+i\psi$ are linear functions of $x+iy$.

Our remarks concerning the relations between Property III and the corresponding classical property will be postponed until after we have given IV.

6. If the conic Γ'_2 corresponding to the point O exists, its curvature at the point O is

$$-\frac{4}{3} \frac{\psi_x \phi^2 + (\psi_y - \phi_x) \phi \psi - \phi_y \psi^2}{\phi^3 [1 + (\psi^2/\phi^2)]^{3/2}}.$$

The curvature at O of the line of force through that point is

$$\frac{\psi_x \phi^2 + (\psi_y - \phi_x) \phi \psi - \phi_y \psi^2}{\phi^3 [1 + (\psi^2/\phi^2)]^{3/2}}.$$

Hence we can state

PROPERTY IV. *If the conic Γ'_2 corresponding to O exists, the ratio of its curvature at O to the curvature (at O) of the line of force through that point is $-4/3$. If the conic does not exist, the curvature of the line of force at O is zero.*

In connection with the converse of this property, we observe that the lines of force are defined geometrically by the property that the tangent at any point has the direction, associated with that point, referred to in the statement of Property II.

The converse of Property IV can be expressed as follows. If a family of curves possessing Properties I to III inclusive also possesses Property IV, the functions ω , μ_1 , μ_2 , and μ_3 , which have appeared above, must satisfy the relation

$$(12) \quad \mu_1 + \mu_2\omega + \mu_3\omega^2 - \omega_x - \omega\omega_y = 0.$$

The relation of Properties III and IV to the classical theory is very much the same as that of Property II. The properties could be taken over, with slight changes of wording, into the classical theory as alternatives to the third and fourth properties in Kasner's set. The properties are not very directly connected with the two given by Kasner, although their converses have the same effect as the converses of his properties in restricting the form of the function $G(x, y, y')$.

7. The parallel asymptotes of the curve Γ'_1 intersect the u -axis in points P_1 and P_2 having the abscissae

$$u = \frac{Ga^2}{2u_0} - 3^{-1/2} \frac{a^2}{A}, \quad u = \frac{Ga^2}{2u_0} + 3^{-1/2} \frac{a^2}{A},$$

respectively. From the point O , as initial point, we draw a vector \overrightarrow{OQ} equal to the vector $\overrightarrow{P_1P_2}$. Then we study the curve Γ_3 described by the terminus Q of this vector as y' varies. The result can be stated as follows:

PROPERTY V. *The curve Γ_3 is a circle which passes through the point O ; and the tangent to the circle at O is perpendicular to the direction, fixed for O , referred to in the statement of Property II.*

For the sake of future use, we note that the equation of the circle Γ_3 is

$$(13) \quad \xi^2 + \eta^2 = \frac{4a^2}{3c^2} (\phi\xi + \psi\eta).$$

Now let us consider the converse of Property V.

If a family of curves possesses Properties I and II, the curve Γ_3 described by the point Q is represented by the equation

$$(14) \quad \xi^2 + \eta^2 = \pm \frac{2^{5/2}}{3} a^2 \xi \left[F \frac{1 + (\eta/\xi)\omega}{(1 + (\eta/\xi)^2)(\omega - (\eta/\xi))} \right]^{1/2},$$

where the symbol F is to be interpreted as $F(x, y, \eta/\xi)$. If the family of curves has also Property V, equation (14) must be of the form

$$(15) \quad \xi^2 + \eta^2 = \frac{2^{5/2}}{3} a^2 \lambda (\xi + \omega \eta),$$

where λ is some function of x and y ; and hence we must have

$$(16) \quad F(x, y, y') = -\lambda^2(1 + y'^2)(1 + \omega y')(y' - \omega).$$

Property V has no analogue in the Newtonian case. This is natural; for we see that the property is connected essentially with the occurrence of the terms $-F$ and $F(1 + Ky''^2)^{1/2}$ in the right-hand member of equation (2), and no such terms exist in the corresponding classical equation.

8. The five properties which we have obtained may be looked upon as the geometrical meaning of the special way in which the derivatives enter into equation (2). They even go somewhat beyond this, in that their converses restrict to some extent the way in which the variables x and y occur in the defining differential equation of a family of curves possessing the properties. However, the most general differential equation defining a family of curves possessing the five properties contains four arbitrary functions of x and y , namely, λ , ω , μ_1 , and μ_3 , whereas equation (2) depends on only two such functions, namely, ϕ and ψ . We must, therefore, proceed to find one or more additional properties to complete the characterization of the families of dynamical trajectories.

9. Referring to equation (13), we see that the ξ -axis intersects the circle Γ_3 in the point M having the coordinates $\xi = (4a^2/3c^2)\phi$, $\eta = 0$. The line through the point O and the center of the circle intersects the circle again in the point M' having the coordinates $\xi = (4a^2/3c^2)\phi$, $\eta = (4a^2/3c^2)\psi$. The distance from the point O to the point M is $OM = (4a^2/3c^2)\phi$, and the distance from the point M to the point M' is $MM' = (4a^2/3c^2)\psi$.

If the conic Γ'_2 corresponding to O exists, the ξ -axis intersects it in the point A having the coordinates $\xi = (3/2)\psi/\psi_x$, $\eta = 0$, and the η -axis intersects it in the point B having the coordinates $\xi = 0$, $\eta = (3/2)\phi/\phi_y$. We let OA and OB , respectively, denote the distances from the point O to the points A and B . Then $OA = (3/2)\psi/\psi_x$, $OB = (3/2)\phi/\phi_y$.

We have immediately

PROPERTY VI. *When the initial point O is changed, the associated circle Γ_3 changes in the manner described by the following equations:*

$$\frac{\partial}{\partial x} MM' = \frac{3}{2} \frac{MM'}{OA} \quad \text{or} \quad 0$$

according as the conic Γ'_2 corresponding to O exists or does not exist;

$$\frac{\partial}{\partial y} OM = \frac{3}{2} \frac{OM}{OB} \quad \text{or} \quad 0$$

according as the conic exists or does not exist.

Conversely, if we take the equation of Γ_3 in the form (15), and the equation of Γ'_2 in the form

$$\omega\xi - \eta = (2/3)(\mu_1\xi^2 + \mu_2\xi\eta + \mu_3\eta^2),$$

and proceed to define distances OA , OB , OM , and MM' as above, we get the results

$$OA = \frac{3}{2} \frac{\omega}{\mu_1}, \quad OB = -\frac{3}{2\mu_3}, \quad OM = \frac{2^{5/2}}{3} a^2 \lambda, \quad MM' = \frac{2^{5/2}}{3} a^2 \lambda \omega.$$

Hence, if a family of curves possessing Properties I, II, III, and V, also possesses Property VI, we have the relations

$$(17) \quad (\lambda\omega)_x = \lambda\mu_1, \quad \lambda_y = -\lambda\mu_3.$$

Since Property VI relates to the circle Γ_3 , it, like Property V, has no analogue in the Newtonian case. On the other hand, having Property VI, we have no need of an analogue of the complicated fifth property in Kasner's set.

10. Now we proceed to show that the six properties which we have obtained are in fact characteristic of the family of relativistic trajectories.

Suppose that a certain three-parameter family of plane curves possesses all six of the properties. Then, as we have seen, the family is defined by a differential equation of the form (2), where the square root has its positive value, and where F , G , H , and K , are given by the formulas (16), (11), (10), and (4), respectively, λ and ω being some functions of x and y , and μ_1 , μ_2 , μ_3 , and ω_1 being defined by the equations (12), (17), and $\omega_1 = \omega$.

Let us define two new functions $\phi(x, y)$ and $\psi(x, y)$ as follows:

$$\lambda = 2^{-1/2} c^{-2} \phi, \quad \omega = \psi/\phi.$$

Then, by (12) and (17), we have the relations

$$\mu_1 = \psi_x/\phi, \quad \mu_2 = (\psi_y - \phi_x)/\phi, \quad \mu_3 = -\phi_y/\phi.$$

When, in the formulas for F , G , and H , we replace λ , ω , μ_1 , μ_2 , and μ_3 by these expressions in terms of ϕ and ψ , we obtain the formulas (3).

Thus, not only does every family of curves defined by a system of equations such as (2) and (3), with the square root positive, possess the six properties given, but also if a three-parameter family of plane curves possesses

the six properties, it is defined by such a system of equations, with suitably chosen functions $\phi(x, y)$ and $\psi(x, y)$. Moreover, if a family of curves is defined by such a system of equations, it is the family of trajectories of a particle moving according to the differential equations of motion (1). Hence, if a family of curves possesses the six properties, it is the family of trajectories of a relativistic particle moving in a suitably chosen positional field of force. Therefore, the set of six properties is characteristic of the families of trajectories of a relativistic particle moving in a plane under forces which are functions (not identically zero) of position only.

It will be observed that the six properties are ordinarily independent, that is, no one of them can be derived from those which precede it.

THE DETERMINATION OF THE FIELD OF FORCE BY THE GEOMETRICAL PROPERTIES OF THE FAMILY OF TRAJECTORIES

11. It is of interest to discuss the way in which a field of force is determined by the geometrical properties of the family of trajectories of a particle moving in the field. In the Newtonian case the geometry of the family of trajectories is incapable of determining more than the direction of the force acting at any point and the ratio of the magnitudes of the forces acting at any two points. On the other hand, in the relativistic case, if the rest-mass of the particle is given,* the geometry of the family of trajectories determines the field of force completely. This is because the right-hand member of equation (2) is not homogeneous and of degree zero in ϕ , ψ , and their partial derivatives, as is the right-hand member of the corresponding classical equation.

When the complete three-parameter family of relativistic trajectories of a particle in a positional field of force is given, we can determine the circle Γ_s corresponding to any point (x, y) , and, by equation (13), this determines the values of the functions ϕ and ψ at (x, y) . The components of the force acting at the point are $m_0\phi(x, y)$ and $m_0\psi(x, y)$. Thus, when the complete three-parameter family of trajectories is given, the field of force is fully determined. However, we are mainly interested in showing that we can determine the force acting at a particular point, or the field of force, without making use of the complete family of trajectories.

12. We shall first show that the force acting at a particular point is determined when three trajectories, passing through that point in the same direction, are given.†

* Throughout this section we suppose that m_0 is given.

† The proof given is based on the assumption that the trajectories are such that two constants, v_2 and v_3^{-1} , are sufficiently small in absolute value. The extent to which this restriction can be removed by the use of continuity considerations has not been investigated.

If the cubic Γ'_1 corresponding to a lineal element (x, y, y') is known, the circle Γ_3 corresponding to the point (x, y) can be constructed immediately;* and then, as has been said above, the force acting at (x, y) is determined. Hence, it will suffice to show that Γ'_1 is determined when three trajectories, passing through (x, y) in the direction determined by y' , are given.

Let T_1, T_2 , and T_3 be three such trajectories. We construct the corresponding three osculating parabolas, determine their foci, and then obtain the inverses of these points with respect to the circle $u^2 + v^2 = a^2$. The coordinates (in the uv -coordinate system) of the last three points will be denoted by $(u_1, v_1), (u_2, v_2), (u_3, v_3)$. We have the equations

$$(18) \quad Ga^2 - 2u_0u_n - 4v_0v_n/3 - \frac{(1 + y'^2)^{3/2}}{4a^4F} v_n [Ga^2 - 2u_0u_n - 2v_0v_n]^2 = 0, \\ n = 1, 2, 3,$$

which we have to solve for F, G , and v_0 , in order to determine Γ'_1 .

From equations (18) we obtain the equations

$$(19) \quad v_1[Ga^2 - 2u_0u_1 - 2v_0v_1]^2[Ga^2 - 2u_0u_n - 4v_0v_n/3] \\ - v_n[Ga^2 - 2u_0u_1 - 4v_0v_1/3][Ga^2 - 2u_0u_n - 2v_0v_n]^2 = 0, \quad n = 2, 3,$$

which we have to solve for G and v_0 . To each solution (G, v_0) of equations (19) there corresponds a unique value of F which is given by any one of equations (18). Now equations (19) have a finite set of solutions. Our problem is to show that only one of these solutions is significant, and to show how the significant solution can be distinguished.

For the time being, let us regard u_1 and v_1 as constants, u_2, v_2, u_3 , and v_3 as variables, and the solutions of equations (19) as pairs of functions of these variables.

It follows from the second part of Property I and the elementary properties of the curve Γ'_1 that the significant solutions of equations (19) are such that as v_2 approaches zero $2u_0u_2$ approaches Ga^2 , and as v_3^{-1} approaches zero u_0r approaches $-v_0$, where $r = u_3/v_3$. Now, for $v_2 = v_3^{-1} = 0$, equations (19) reduce to

$$v_1[Ga^2 - 2u_0u_1 - 2v_0v_1]^2[Ga^2 - 2u_0u_2] = 0, \\ [Ga^2 - 2u_0u_1 - 4v_0v_1/3][2u_0r + 2v_0]^2 = 0.$$

Hence, two of the solutions of (19) satisfy the above elementary criteria for

* The construction is an easy consequence of Properties II and V and the definition of Γ_3 . There is an ambiguity in the construction, arising from the two possible ways of drawing a vector from one of the intersections of asymptotes of Γ'_1 to the other. However, this ambiguity is removed when we take account of the fact that a trajectory lies on that side of its tangent toward which the force is directed.

significance, and we must seek an additional criterion to distinguish between them.

A little consideration of the properties of the curve Γ'_1 suffices to show that if the absolute value of v_3 is large, we have a relation of the form

$$v_0 = -u_0 r + (Ga^2 + C)/(2v_3) + O(v_3^{-2}),$$

where C is a constant which has the same sign as the product $v_0 v_3$. On the other hand, if we regard the second of equations (19) as a relation between an independent variable v_3 and a dependent variable v_0 , we readily find that the two roots which reduce to $-u_0 r$ for $v_3^{-1} = 0$ are given, for small values of v_3^{-1} , by the expansions

$$(20) \quad v_0 = -u_0 r + \frac{Ga^2}{2v_3} \pm \frac{1}{2v_3} (Ga^2 - 2u_0 u_1 + 2u_0 v_1 r) \cdot \left[\frac{-2u_0 v_1 r/3}{Ga^2 - 2u_0 u_1 + 4u_0 v_1 r/3} \right]^{1/2} + O(v_3^{-2}).$$

Hence, only that solution of equations (19) is significant which, for small values of v_3^{-1} , gives the third term in the right-hand member of (20) the same sign as $-u_0 r$. (It is easily shown that the square root in (20) is real when v_3^{-1} is small.)

To summarize: When the three trajectories T_1 , T_2 , and T_3 , are given, equations (18) determine a finite set of solutions (F, G, v_0). Only one of these solutions is significant, namely, the one which behaves as described above when v_2 and v_3^{-1} , regarded momentarily as variables, approach zero. When the significant solution has been obtained, the curve Γ'_1 corresponding to the lineal element (x, y, y') is determined, and the force acting at (x, y) can be calculated.

13. We shall show that, subject to certain restrictions of an analytical character, a positional field of force is determined throughout a neighborhood of a point when the force acting at the point and four one-parameter families of trajectories, each of which covers the neighborhood simply,* are given.

Let us suppose that in a neighborhood of a point P we have four one-parameter families of curves, of the type just described, which are known to be trajectories of a particle of rest-mass m_0 in an unknown positional field of force. We also suppose that the force acting at the point P is known.

The equations of the curves of the four given families will be written

$$(21) \quad f_n(x, y) = a_n, \quad n = 1, 2, 3, 4,$$

* That is, so that through each point of the neighborhood there passes just one curve of each family.

where the a 's are the parameters of the families. We assume that the left-hand members of these equations are analytic functions of their arguments, and that none of the partial derivatives $\partial f_n / \partial y$ vanish in the neighborhood of P .

The functions $y_n = y_n(x, a_n)$ defined by equations (21) satisfy the relations

$$(22) \quad \begin{aligned} (\psi - \phi y_n') y_n''' &= -\frac{1}{2c^4} (1 + y_n'^2) (\psi - \phi y_n')^2 (\phi + \psi y_n') \\ &+ [\psi_x + (\psi_y - \phi_x) y_n' - \phi_y y_n'^2] y_n'' - 3\phi y_n'^2 \\ &+ \frac{1}{2c^4} (1 + y_n'^2) (\psi - \phi y_n')^2 (\phi + \psi y_n') \\ &\cdot \left[1 + \frac{4c^4}{(1 + y_n'^2)^2 (\psi - \phi y_n')^2} y_n'^2 \right]^{1/2}, \end{aligned}$$

where $m_0\phi(x, y)$ and $m_0\psi(x, y)$ are the (unknown) components of the force acting at (x, y) , and where y_n' , y_n'' , and y_n''' are to be interpreted, in an obvious way, as definite analytic functions of x, y determined by equations (21).

We assume that the determinant

$$\begin{vmatrix} y_1' & y_1' y_1'' & y_1'^2 y_1''' & y_1'^2 \\ y_2' & y_2' y_2'' & y_2'^2 y_2''' & y_2'^2 \\ y_3' & y_3' y_3'' & y_3'^2 y_3''' & y_3'^2 \\ y_4' & y_4' y_4'' & y_4'^2 y_4''' & y_4'^2 \end{vmatrix}$$

does not vanish at the point P . (This determinant cannot vanish at every point of a neighborhood of P , for all choices of the one-parameter families of trajectories, unless the force vanishes throughout the neighborhood. Otherwise, the family of all trajectories in a nonzero field of force would consist merely of a finite set of two-parameter families.) Consequently, we can solve equations (22) algebraically for ψ_x , $(\psi_y - \phi_x)$, ϕ_y , and the ϕ which appears in the third terms of the right-hand members, obtaining a set of relations which we shall write schematically as follows:

$$(23) \quad \begin{aligned} \psi_x &= f(x, y, \phi, \psi), & \psi_y - \phi_x &= g(x, y, \phi, \psi), \\ \phi_y &= h(x, y, \phi, \psi), & \phi &= k(x, y, \phi, \psi). \end{aligned}$$

It is to be emphasized that, in virtue of the given equations (21), the right-hand members of equations (23) are entirely definite analytic functions of the arguments indicated.

It follows from equations (23) that we have the system of partial differential equations

$$(24) \quad \begin{aligned} \phi_x &= (k_x + k_\psi f)/(1 - k_\phi), & \phi_y &= h, \\ \psi_x &= f, & \psi_y &= g + (k_x + k_\psi f)/(1 - k_\phi). \end{aligned}$$

By our assumption that the given curves (21) are known to be trajectories in an unknown positional field of force, and that the force acting at P is given, the system of equations (24) is satisfied by a pair of functions $\phi(x, y)$, $\psi(x, y)$, which have given values at the point P .

If the two equations forming the conditions for integrability of the system (24) are satisfied identically, the system is completely integrable. In this case the field of force is determined throughout a neighborhood of P by the differential equations (24) and the given value of the force acting at P , at least if the coordinates of P and the values of ϕ and ψ at P form a system of values in the neighborhood of which the right-hand members of the equations are holomorphic. If the conditions for integrability are not satisfied identically, and are two independent equations, these equations determine implicitly a certain finite number of distinct pairs of functions $\phi(x, y)$ and $\psi(x, y)$. Then, if the point P is one at which the distinct pairs of functions have distinct pairs of values, the field of force is determined throughout a neighborhood of P by the conditions for integrability and the given values of ϕ and ψ at P . There is a third conceivable case, namely, that in which just one of the two conditions for integrability is not satisfied identically, or in which, while neither condition is satisfied identically, the two conditions are equivalent to a single equation. In this case we have in effect to deal with a completely integrable system of partial differential equations in one of the unknown functions (say ϕ) and an equation which determines the other unknown function (say ψ) implicitly in terms of x, y , and ϕ . Again we see that if the point P is such that certain conditions of analyticity are satisfied, and certain distinct pairs of functions have distinct pairs of values, the field of force is determined throughout a neighborhood of P by the equations (24), the conditions for integrability, and the given value of the force acting at P .

THE POINT TRANSFORMATIONS WHICH CONVERT EVERY FAMILY OF DYNAMICAL TRAJECTORIES INTO A FAMILY OF DYNAMICAL TRAJECTORIES

14. Kasner has shown that, in the Newtonian case, collineations are the only point transformations of the plane which convert every three-parameter family of trajectories (belonging to a positional field of force) into such a family of curves.* We proceed now to obtain the corresponding result for the relativistic case.

* In general, the fields of force corresponding to the original family and the transformed family, respectively, are different.

Suppose that we have a family of dynamical trajectories, defined by a system of equations such as (2), (3). We apply a point transformation

$$(25) \quad x = x(\bar{x}, \bar{y}), \quad y = y(\bar{x}, \bar{y}),$$

where the functions $x(\bar{x}, \bar{y})$, $y(\bar{x}, \bar{y})$ are of class C^2 , and the Jacobian $x_{\bar{x}}y_{\bar{y}} - x_{\bar{y}}y_{\bar{x}}$ does not vanish in the region under consideration; and we require that the transformation be specialized so that the transformed family of curves shall be defined by a system of equations of the form

$$\begin{aligned} \bar{y}''' &= -\bar{F} + \bar{G}\bar{y}'' + \bar{H}\bar{y}'^2 + \bar{F}(1 + \bar{K}\bar{y}'^2)^{1/2}, \\ \bar{F} &= \frac{1}{2c^4} (1 + \bar{y}'^2)(\bar{\psi} - \bar{y}'\bar{\phi})(\bar{\phi} + \bar{y}'\bar{\psi}), \quad \bar{H} = -\frac{3\bar{\phi}}{\bar{\psi} - \bar{\phi}\bar{y}'}, \\ \bar{G} &= \frac{\bar{\psi}_{\bar{x}} + (\bar{\psi}_{\bar{y}} - \bar{\phi}_{\bar{x}})\bar{y}' - \bar{\phi}_{\bar{y}}\bar{y}'^2}{\bar{\psi} - \bar{\phi}\bar{y}'}, \quad \bar{K} = 4c^4(1 + \bar{y}'^2)^{-2}(\bar{\psi} - \bar{\phi}\bar{y}')^{-2}. \end{aligned}$$

Here $\bar{\phi}$ and $\bar{\psi}$ denote functions of \bar{x} and \bar{y} , and the primes denote differentiation with respect to \bar{x} .

On transforming equations (2), (3) by means of (25) and its extensions, we obtain an equation of the form

$$(26) \quad \bar{y}''' = R_1 + R_2\bar{y}'' + R_3\bar{y}'^2 + R_4 \left[1 + \frac{4c^4(\alpha + \beta\bar{y}' + \gamma\bar{y}'^2 + \delta\bar{y}'^3 + \epsilon\bar{y}'^4)}{[(x_{\bar{x}} + y_{\bar{y}}\bar{y}')^2 + (y_{\bar{x}} + x_{\bar{y}}\bar{y}')^2]^2 [A(x_{\bar{x}} + x_{\bar{y}}\bar{y}') - B(y_{\bar{x}} + y_{\bar{y}}\bar{y}')]^2} \right]^{1/2},$$

where A and B are functions of \bar{x} and \bar{y} , the R 's are functions of \bar{x} , \bar{y} , and \bar{y}' , which are rational in \bar{y}' , and where

$$(27) \quad \begin{aligned} \alpha &= x_{\bar{x}}y_{\bar{x}\bar{x}} - x_{\bar{x}\bar{x}}y_{\bar{x}}, & \beta &= 2(x_{\bar{x}}y_{\bar{x}\bar{y}} - x_{\bar{x}\bar{y}}y_{\bar{x}}) + (x_{\bar{y}}y_{\bar{x}\bar{x}} - x_{\bar{x}\bar{x}}y_{\bar{y}}), \\ \gamma &= 2(x_{\bar{y}}y_{\bar{x}\bar{y}} - x_{\bar{x}\bar{y}}y_{\bar{y}}) + (x_{\bar{x}}y_{\bar{y}\bar{y}} - x_{\bar{y}\bar{y}}y_{\bar{x}}), & \delta &= x_{\bar{y}}y_{\bar{y}\bar{y}} - x_{\bar{y}\bar{y}}y_{\bar{y}}, \\ \epsilon &= x_{\bar{x}}y_{\bar{y}} - x_{\bar{y}}y_{\bar{x}}. \end{aligned}$$

In order that equation (26) shall reduce to the required form, for all choices of the functions $\phi(x, y)$, $\psi(x, y)$ in the original equations, it is obviously necessary that we have the relations

$$(28) \quad \alpha = \beta = \gamma = \delta = 0, \quad x_{\bar{x}}^2 + y_{\bar{x}}^2 = x_{\bar{y}}^2 + y_{\bar{y}}^2, \quad x_{\bar{x}}x_{\bar{y}} + y_{\bar{x}}y_{\bar{y}} = 0.$$

It readily follows from equations (27) and (28) that $x_{\bar{x}}$, $x_{\bar{y}}$, $y_{\bar{x}}$, and $y_{\bar{y}}$ must be constants, and that we must have the relations

$$y_{\bar{x}} = \pm x_{\bar{y}}, \quad y_{\bar{y}} = \mp x_{\bar{x}}.$$

Hence it is necessary, in order that a point transformation shall convert every three-parameter family of relativistic trajectories into a family of rela-

tivistic trajectories, that the transformation be a rigid motion, a magnification, a reflection with respect to a straight line, or a combination of such transformations. Also, we easily see that this condition is sufficient to insure that the transformation has the required property.

NATURAL FAMILIES OF RELATIVISTIC TRAJECTORIES

15. So far we have been considering the family of all possible trajectories of a particle in a positional field of force. Now we wish to study certain important subfamilies of trajectories, which we call natural families. Since, in the study of natural families, we can easily deal with the case of a particle moving in three-dimensional space, we shall do so. The results for the case in which the particle moves in a fixed plane can be obtained by a simple specialization.

Let us consider a relativistic particle, of rest-mass m_0 , moving in three-dimensional space under a force which is derived from a potential energy function $V(x, y, z)$. Here x, y , and z are the rectangular coordinates of the particle with respect to a fixed set of axes. The differential equations of motion are the following:

$$(29) \quad \begin{aligned} \frac{d}{dt} \left[\dot{x} \left(1 - \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{c^2} \right)^{-1/2} \right] &= -\phi_x, \\ \frac{d}{dt} \left[\dot{y} \left(1 - \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{c^2} \right)^{-1/2} \right] &= -\phi_y, \\ \frac{d}{dt} \left[\dot{z} \left(1 - \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{c^2} \right)^{-1/2} \right] &= -\phi_z, \end{aligned}$$

where $\phi = V/m_0$.

Equations (29) possess the integral

$$(30) \quad c^2 \left(1 - \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{c^2} \right)^{-1/2} = h - \phi,$$

where h is a constant of integration. Hence, the five-parameter family of trajectories defined by equations (29) consists of ∞^1 four-parameter families, each particular one of which corresponds to a particular value of h . Each of these four-parameter families will be called a natural family of trajectories.*

For the sake of convenience, we shall write $h - \phi = \Phi$.

It is easily shown that the defining differential equations of the natural family of trajectories corresponding to h can be written in the form

* Natural families of trajectories of classical particles are defined in an analogous way. See Kasner, *Differential-Geometric Aspects of Dynamics*, p. 34.

$$(31) \quad \begin{aligned} (1 + y'^2 + z'^2)^{-1/2} \frac{d}{dx} \left[y' \left(\frac{\Phi^2 - c^4}{1 + y'^2 + z'^2} \right)^{1/2} \right] &= \frac{\partial}{\partial y} (\Phi^2 - c^4)^{1/2}, \\ (1 + y'^2 + z'^2)^{-1/2} \frac{d}{dx} \left[z' \left(\frac{\Phi^2 - c^4}{1 + y'^2 + z'^2} \right)^{1/2} \right] &= \frac{\partial}{\partial z} (\Phi^2 - c^4)^{1/2}, \end{aligned}$$

where $y' = dy/dx$, $z' = dz/dx$.

Now in Newtonian dynamics the equations corresponding to (29) are $\ddot{x} = -\phi_x$, $\ddot{y} = -\phi_y$, $\ddot{z} = -\phi_z$; the equation corresponding to (30) is $(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)/2 = h - \phi$; and hence the differential equations defining the natural family of trajectories corresponding to h are

$$(32) \quad \begin{aligned} (1 + y'^2 + z'^2)^{-1/2} \frac{d}{dx} \left[y' \left(\frac{2\Phi}{1 + y'^2 + z'^2} \right)^{1/2} \right] &= \frac{\partial}{\partial y} (2\Phi)^{1/2}, \\ (1 + y'^2 + z'^2)^{-1/2} \frac{d}{dx} \left[z' \left(\frac{2\Phi}{1 + y'^2 + z'^2} \right)^{1/2} \right] &= \frac{\partial}{\partial z} (2\Phi)^{1/2}. \end{aligned}$$

On comparing the systems of equations (31) and (32), we get the

THEOREM. *If the constants E_1 , E_2 , A , and m_0 , and the functions $V_1(x, y, z)$ and $V_2(x, y, z)$ are such that we have identically*

$$A[E_1 - V_1(x, y, z)] = [E_2 - V_2(x, y, z)]^2 - m_0^2 c^4,$$

the natural family of trajectories of a classical particle moving with (classical) total energy E_1 in the field of force derived from the potential energy function $V_1(x, y, z)$ is identical with the natural family of trajectories of a relativistic particle, of rest-mass m_0 , moving with (relativistic) total energy E_2 in the field of force derived from the potential energy function $V_2(x, y, z)$.

Undoubtedly, the content of this theorem is more or less familiar, since it is an immediate consequence of the well known fact that, whereas the classical trajectories are defined by the principle of least action

$$\delta \int (E_1 - V_1)^{1/2} ds = 0,$$

the relativistic trajectories are defined by the principle

$$\delta \int [(E_2 - V_2)^2 - m_0^2 c^4]^{1/2} ds = 0.$$

However, the theorem does not seem to be stated explicitly in any of the readily accessible literature.

We have seen in the preceding sections that the sets of properties which

characterize the families of all trajectories of a particle in an arbitrary field of force are very different in the classical and relativistic cases, respectively. The present theorem shows that if we consider not the families of all trajectories but only natural families of trajectories, the characteristic properties are the same in the two cases. The characteristic properties of a natural family of trajectories have been given by Kasner.*

* *Differential-Geometric Aspects of Dynamics*, pp. 37-42. See also J. Lipka, these Transactions vol. 13 (1912), pp. 77-95.

BELL TELEPHONE LABORATORIES,
NEW YORK, N. Y.

THE DIFFERENTIAL GEOMETRY OF SERIES OF LINEAL ELEMENTS*

BY

JOHN DE CICCIO

1. **Introduction.** We shall begin by considering certain simple operations or transformations on the oriented lineal elements of the plane. A *turn* T_α converts each element into one having the same point and making a fixed angle α with the original direction. By a *slide* S_k , the line of the element remains the same and the point moves along the line a fixed distance k . These transformations together generate a continuous group of three parameters which we call the *whirl group* W_3 . The group of whirls W_3 is isomorphic to the group of rigid motions M_3 . These two three-parameter groups are commutative and together form a new group of six parameters which we term the *whirl-motion group* G_6 . In preceding papers (see the bibliography at the end of this paper), Kasner and the author developed the geometry of this group G_6 . In this paper, we wish to give the differential geometry of the series of lineal elements in the plane with respect to the whirl-motion group G_6 .

A set of ∞^1 elements is called a *series*: this includes a union (curve or point) as a special case. A collection of ∞^2 elements is termed a *field*, which of course corresponds to a differential equation of first order, $F(x, y, y') = 0$. The totality of ∞^3 elements of the plane is called the *opulence*.

A *turbine* is the series which is obtained by applying a turn T_α to each element of an oriented circle (the outer circle). It is said to be *nonlinear* or *linear* according as the base circle is not or is a straight line. A *nonlinear flat field* consists of the ∞^2 elements cocircular with a given element, called the *center* or *central element*. A *linear flat field* is the set of ∞^2 elements on the ∞^1 oriented lines, which are parallel and possess the same orientation.

In this paper, we shall consider the tangent turbines and the osculating flat fields of a series of lineal elements. We shall find the necessary and sufficient conditions that ∞^1 limaçon (circular) series be the osculating limaçon (circular) series of a general (equiparallel) series (Theorems 11 and 16). We shall define the *curvature* and *torsion* of any series (formulas (38), (39), and (47)). The curvature and torsion of a series \bar{S} conjugate to a given series S will be obtained in terms of the curvature and torsion of the given series S . Finally we shall find that *any two general (equiparallel) series, which have their curvatures and torsions the same functions of the angle u (arc*

* Presented to the Society, March 26, 1937; received by the editors February 16, 1939.

length s), are equivalent under the whirl-motion group G_6 (Theorems 20 and 21). This then gives us the intrinsic equations of a series of lineal elements in the geometry of the whirl-motion group G_6 .

For the analytic representation, it will be convenient to define an element by the hessian coordinates (u, v, w) where v is the length of the perpendicular from the origin, u is the angle between the perpendicular and the initial line, and w is the distance between the foot of the perpendicular and the point of the element.

2. The tangent turbines of a general series. Any series which consists of ∞^1 nonparallel elements is termed a *general series*, whereas an *equiparallel series* consists of ∞^1 parallel elements. Thus a general series is never contained in a linear flat field, while an equiparallel series always lies in a linear flat field.

Any general series is given by the equations

$$(1) \quad v = v(u), \quad w = w(u),$$

while any equiparallel series is given by the equations

$$(2) \quad u = c, \quad w = w(v),$$

where c is a constant.

The points of the elements of a series form a union which we call the *point-union* of the series. The lines of the elements of a general series are the tangent lines of a union which is called the *line-union* of the general series. For an equiparallel series, there is no line-union since the lines of the element all have a common direction. The point-union is called the *base curve* of the equiparallel series.

A nonlinear turbine is a general series. Its point-union is a circle, called the *outer circle*, and its line-union is also a circle, called the *inner circle*. These two circles are concentric, and their common center is called the center of the turbine. Of course, the inner circle is in the interior of the outer circle.

From the preceding remarks, we may have the following construction for a nonlinear turbine in addition to the one given in §1. A nonlinear turbine is the series which is obtained by applying a slide S_s to each element of an oriented circle (the inner circle). From this, we find that the equations of a nonlinear turbine are

$$(3) \quad v = a \cos u + b \sin u + r, \quad w = -a \sin u + b \cos u + s,$$

where (a, b) are the cartesian coordinates of the center, r is the radius of the inner circle, and s is the constant distance of the slide S_s . We call (a, b, r, s) a set of nonlinear turbine coordinates.

From (3) or by synthetic reasoning, it may be shown that (1) two elements which are not simultaneously parallel and of the same orientation determine a unique nonlinear turbine, and (2) two nonlinear turbines possess either one common element or no common elements.

If a one-parameter family of series has the property that consecutive series have a common element, then the family is called a *set of enveloping series*. The locus of intersection of consecutive series of the family is called the *envelope* of the family. Thus the one-parameter family of general series $v=v(u, t)$, $w=w(u, t)$ is a set of enveloping series if and only if the equations $v_t=0$ and $w_t=0$ have a common solution in u . The envelope is then given by the two eliminants with respect to t of these four equations.

Let two series S_1 and S_2 possess a common element E_0 . These two series are said to be *tangent* (or to have contact of first order) at E_0 if and only if they have two consecutive elements in common at E_0 . Thus the two general series $S_1: v=v_1(u)$, $w=w_1(u)$, and $S_2: v=v_2(u)$, $w=w_2(u)$ are tangent at the common element $E_0(u_0, v_0, w_0)$ if and only if

$$(4) \quad \begin{aligned} v_0 &= v_1(u_0) = v_2(u_0), & w_0 &= w_1(u_0) = w_2(u_0), \\ v'_1(u_0) &= v'_2(u_0), & w'_1(u_0) &= w'_2(u_0). \end{aligned}$$

Let $S_t: v=v(u, t)$, $w=w(u, t)$ denote a one-parameter family of enveloping series, and let S denote the envelope of this family. From the equations of the envelope S and by (4), it easily follows that *any series S_t of the one-parameter family of enveloping series is tangent to the envelope S at any one of their common elements*.

If a one-parameter family of turbines is an enveloping set of turbines, then we shall say that the turbines are the *tangent turbines* of the envelope.

THEOREM 1. *The ∞^1 nonlinear turbines*

$$(5) \quad v = a(t) \cos u + b(t) \sin u + r(t), \quad w = -a(t) \sin u + b(t) \cos u + s(t)$$

constitute a set of tangent turbines if and only if

$$(6) \quad a'^2 + b'^2 = r'^2 + s'^2.$$

For, this is the condition that the equations

$$(7) \quad a' \cos u + b' \sin u + r' = 0, \quad -a' \sin u + b' \cos u + s' = 0$$

be compatible in u .

The envelope of the ∞^1 nonlinear turbines is given by the equations (5) and (7). Solving (7) for $\cos u$ and $\sin u$, we obtain the

COROLLARY. *The series to which the nonlinear turbines of Theorem 1 are the*

tangent turbines either consists of one element or is a general series. It is given by the equations

$$(8) \quad \cos u = \frac{-a'r' - b's'}{a'^2 + b'^2}, \quad \sin u = \frac{a's' - b'r'}{a'^2 + b'^2},$$

$$v = a \cos u + b \sin u + r, \quad w = -a \sin u + b \cos u + s.$$

The envelope (8) of the tangent turbines is given by the equations (5), where the value of t in terms of u is defined by the equations (7). If equations (5), subject to the conditions (7), are differentiated totally with respect to u , the resulting equations are

$$(9) \quad v' = -a \sin u + b \cos u, \quad w' = -a \cos u - b \sin u,$$

where the accent denotes total differentiation with respect to u . But these equations and (5) may be solved for a, b, r, s . Thus, we have established the following result.

THEOREM 2. *The tangent turbines of the general series (1) are the nonlinear turbines whose parameter values are*

$$(10) \quad a = -v' \sin u - w' \cos u, \quad b = v' \cos u - w' \sin u,$$

$$r = v + w', \quad s = -v' + w,$$

where the accent denotes differentiation with respect to u .

It is noted that, if a general series is a curve, then the tangent turbines are the osculating circles of the curve.

A tangent turbine of a general series S at an element E may be defined as the unique limiting turbine of the set of nonlinear turbines such that any nonlinear turbine of this set contains the element E and any other nearby element of S .

3. The tangent turbines of an equiparallel series. A linear turbine is the series which is obtained by applying a turn T_ω to each element of an oriented straight line. Thus a linear turbine is an equiparallel series whose base curve is a straight line. The equations of a linear turbine are

$$(11) \quad u = U - \omega, \quad v \cos \omega + w \sin \omega = V,$$

where (U, V) are the hessian coordinates of the base line and ω is the constant angle of the turn T_ω .

By the same process of reasoning as that used in the preceding section, we obtain the following results.

THEOREM 3. *The ∞^1 linear turbines*

$$(12) \quad u = U(t) - \omega(t), \quad v \cos \omega(t) + w \sin \omega(t) = V(t)$$

constitute a set of tangent turbines if and only if

$$(13) \quad U' = \omega' \neq 0.$$

COROLLARY. *The series to which the linear turbines of Theorem 3 are the tangent turbines either consists of one element or is an equiparallel series. It is given by the equations*

$$(14) \quad u = U - \omega = \text{const.}, \quad v = V \cos \omega - \frac{V'}{\omega'} \sin \omega, \quad w = V \sin \omega + \frac{V'}{\omega'} \cos \omega.$$

THEOREM 4. *The tangent turbines of the equiparallel series (2) are the linear turbines whose parameter values are*

$$(15) \quad U = c - \arctan \frac{1}{w'} + n\pi, \quad V = \pm \frac{vw' - w}{(1 + w'^2)^{1/2}}, \quad \omega = -\arctan \frac{1}{w'} + n\pi,$$

where the accent denotes differentiation with respect to v .

A tangent turbine of an equiparallel series S at an element E may be defined as the unique limiting turbine of the set of linear turbines such that any linear turbine of the set contains the element E and any other nearby element of S .

It may be now observed that two series at a common element E are tangent at E if and only if they have the same tangent turbine at E .

4. Conjugate series of elements. Two turbines T and \bar{T} are said to be conjugate if they have the same circle as point locus and the elements of the two turbines are symmetrically related to the elements of the circle.

Two series S and \bar{S} are said to be conjugate if there exists a one-to-one correspondence between their elements in such a way that the tangent turbines of the two series at the corresponding elements are conjugate turbines.

By Theorem 1, we find

THEOREM 5. *For any general series S , there always exists one and only one conjugate series \bar{S} which either consists of one element or is a general series. This series \bar{S} is given by the equations*

$$(16) \quad \cos \bar{u} = \frac{-a'r' + b's'}{a'^2 + b'^2}, \quad \sin \bar{u} = \frac{-a's' - b'r'}{a'^2 + b'^2},$$

$$\bar{v} = a \cos \bar{u} + b \sin \bar{u} + r, \quad \bar{w} = -a \sin \bar{u} + b \cos \bar{u} - s,$$

where (a, b, r, s) are the parameter values of the tangent turbines of S .

It is noted that the only self-conjugate series are the unions.

It may be observed that, if the conjugate series of a general series consists

of only one element \bar{E} , then S is contained in the nonlinear flat field whose central element is \bar{E} . In this case, we shall say that S is a *co-flat series*.

Obviously if an equiparallel series is not a turbine, then there is no series which is conjugate to it.

5. The osculating flat fields of a series of elements. The flat field which has three consecutive elements in common with a series S at an element E of S is called the *osculating flat field* of the series S at the element E .

THEOREM 6. *The osculating flat fields of a general series S are the nonlinear flat fields whose central elements are the elements of the series \bar{S} conjugate to S .*

If S is a co-flat series, then S has one and only one osculating flat field, namely, the nonlinear flat field in which it is contained.

THEOREM 7. *Any equiparallel series has one and only one osculating flat field, namely, the linear flat field in which it is contained.*

Of course, the tangent turbine of a series S at an element E of S is contained in the osculating flat field of S at E .

6. The limaçon series. Let T be a nonlinear turbine (not a point-turbine), let \bar{E} be a fixed element on the conjugate turbine \bar{T} of T , and let γ be a real number. Let O be the point of \bar{E} , and let P be the point of any element E of the turbine T . On the line (OP) , let us select the points P_i , ($i=1, 2$), such that $d(P, P_i)=2\gamma$. Let E_i be the element whose point is P_i and whose direction is that of E . By this construction, to each element E of T , there are associated two elements E_1 and E_2 . The totality of elements E_1, E_2 is called a *limaçon series* with central turbine T and radius γ .

Let T be a point-turbine (point, or star), let \bar{E} be a fixed element of T , and let γ be a real number. Let L be the angle bisector of the angle whose initial and terminal sides are the lines of \bar{E} and of any element E of T respectively. On L , let us select the points P_i , ($i=1, 2$), such that $d(O, P_i)=2\gamma$, where O is the point of T . Let E_i be the element whose point is P_i and whose direction is that of E . By this construction, to each element E of T , there are associated two elements E_1 and E_2 . The totality of elements E_1, E_2 is called a *limaçon series* with central turbine T and radius γ .

The equations of any limaçon series are

$$(17) \quad \begin{aligned} v &= A \cos u + B \sin u + 2\gamma \sin(u - \bar{u})/2 + R, \\ w &= -A \sin u + B \cos u + 2\gamma \cos(u - \bar{u})/2 + S, \end{aligned}$$

where (A, B, R, S) are the parameters of the central turbine T , \bar{u} is the normal angle of the fixed element \bar{E} , and γ is the radius of the limaçon series.

Upon setting

$$(18) \quad C = -2\gamma \sin \bar{u}/2, \quad D = 2\gamma \cos \bar{u}/2,$$

the equations (17) of the limaçon series take the form

$$(19) \quad \begin{aligned} v &= A \cos u + B \sin u + C \cos u/2 + D \sin u/2 + R, \\ w &= -A \sin u + B \cos u - C \sin u/2 + D \cos u/2 + S. \end{aligned}$$

We call $L(A, B, C, D, R, S)$ a set of limaçon series coordinates. Obviously,

$$L(A, B, C, D, R, S) = L(A, B, -C, -D, R, S).$$

The point-union of the limaçon series (19) is the limaçon

$$(20) \quad X + iY = (A + iB) + (C + iD)e^{iu/2} + (R + iS)e^{iu},$$

while the line-union is

$$(21) \quad X + iY = (A + iB) + (1/4)(C - iD)e^{3iu/2} + (3/4)(C + iD)e^{iu/2} + Re^{iu}.$$

From (10) and (19), we obtain

THEOREM 8. *A limaçon series is a co-flat series. The tangent turbines of a limaçon series are such that their conjugate turbines contain the element \bar{E} and such that their centers are on the circle with center (A, B) and radius γ .*

From this theorem, we derive

THEOREM 9. *Three co-flat nonlinear turbines which do not all contain one element determine a unique limaçon series. Three elements, no two of which are parallel, and which do not all lie on one turbine, determine four limaçon series. Three elements, two of which are parallel without all being parallel, determine two limaçon series.*

7. The circular series. The equiparallel series whose point-union is a circle with center (A, B) and radius γ is called the *circular series* with center (A, B) and radius γ .

The equations of a circular series are

$$(22) \quad u = c, \quad (v - \alpha)^2 + (w - \beta)^2 = \gamma^2,$$

where (c, α, β) are the hessian coordinates of the element whose point is (A, B) and whose inclination is $c + \pi/2$. Thus we must have the relations

$$(23) \quad A = \alpha \cos c - \beta \sin c, \quad B = \alpha \sin c + \beta \cos c.$$

THEOREM 10. *Three parallel elements which are not all on one turbine determine a unique circular series. Three linear turbines which all possess the same common direction and no two of whose base lines are parallel determine four circular series. Three linear turbines which all possess the same common direction and only two of whose base lines are parallel determine two circular series.*

8. **The osculating limaçon series of a general series.** Let two series S_1 and S_2 possess a common element E_0 . These two series are said to be *osculating* (or to have contact of the second order) at E_0 if and only if they have three consecutive elements in common at E_0 . Thus the two general series $S_1: v=v_1(u), w=w_1(u)$, and $S_2: v=v_2(u), w=w_2(u)$ are said to be osculating at the common element $E_0(u_0, v_0, w_0)$ if and only if

$$(24) \quad \begin{aligned} v_0 &= v_1(u_0) = v_2(u_0), & w_0 &= w_1(u_0) = w_2(u_0), \\ v'_1(u_0) &= v'_2(u_0), & w'_1(u_0) &= w'_2(u_0), \\ v''_1(u_0) &= v''_2(u_0), & w''_1(u_0) &= w''_2(u_0). \end{aligned}$$

Let us consider the one-parameter family of enveloping series $S_t: v=v(u, t), w=w(u, t)$. Every series S_t of the family is tangent to the envelope S . If every series S_t of the family is also an osculating series of the envelope S , then the family is called a *set of osculating series*. Our given one-parameter family of series is a set of osculating series if and only if the four equations

$$(25) \quad v_t = 0, \quad w_t = 0, \quad v_{ut} = 0, \quad w_{ut} = 0$$

have a common solution in u . The series S_t of the family are then the osculating series of the envelope S .

THEOREM 11. *The ∞^1 limaçon series*

$$(26) \quad \begin{aligned} v &= A(t) \cos u + B(t) \sin u + C(t) \cos u/2 + D(t) \sin u/2 + R(t), \\ w &= -A(t) \sin u + B(t) \cos u - C(t) \sin u/2 + D(t) \cos u/2 + S(t) \end{aligned}$$

constitute a set of osculating limaçon series if and only if

$$(27) \quad \begin{aligned} 4(A'R' - B'S') &= C'^2 - D'^2, & 2(A'S' + B'R') &= C'D', \\ A'^2 + B'^2 &= R'^2 + S'^2. \end{aligned}$$

For, these are the conditions that the equations

$$(28) \quad \begin{aligned} A' \cos u + B' \sin u - R' &= 0, & -A' \sin u + B' \cos u - S' &= 0, \\ C' \cos u/2 + D' \sin u/2 + 2R' &= 0, & -C' \sin u/2 + D' \cos u/2 + 2S' &= 0, \end{aligned}$$

which are equivalent to the equations (25) for the ∞^1 limaçon series (26), be compatible in u .

The envelope of the ∞^1 limaçon series is given by the equations (26) and (28). Solving (28) for $\cos u$ and $\sin u$, we obtain the

COROLLARY. *The series to which the ∞^1 limaçon series of Theorem 11 are the osculating limaçon series either consists of one element or is a general series. It is given by the equations*

$$\begin{aligned}
 (29) \quad \cos u &= \frac{A'R' + B'S'}{A'^2 + B'^2}, & \sin u &= \frac{-A'S' + B'R'}{A'^2 + B'^2}, \\
 v &= A \cos u + B \sin u + C \cos u/2 + D \sin u/2 + R, \\
 w &= -A \sin u + B \cos u - C \sin u/2 + D \cos u/2 + S.
 \end{aligned}$$

The series S of (29) to which the limaçon series are the osculating limaçon series is given by the equations (26), where the value of t in terms of u is defined by equations (28). If equations (26), subject to the conditions (28), are differentiated totally with respect to u and if the results are again differentiated totally with respect to u , we find that these are equivalent to

$$\begin{aligned}
 (30) \quad C \cos u/2 + D \sin u/2 &= -4s', & -C \sin u/2 + D \cos u/2 &= 4r', \\
 A \cos u + B \sin u &= -w' + 2s', & -A \sin u + B \cos u &= v' - 2r',
 \end{aligned}$$

where r and s are the last two parameters of the tangent turbines to the series S . Solving (26) and (30) for A, B, C, D, R, S , we obtain the

THEOREM 12. *The osculating limaçon series of the general series S of (1) are those whose parameter values are*

$$\begin{aligned}
 (31) \quad A &= a + 2r' \sin u + 2s' \cos u, & B &= b - 2r' \cos u + 2s' \sin u, \\
 C &= -4r' \sin u/2 - 4s' \cos u/2, & D &= 4r' \cos u/2 - 4s' \sin u/2, \\
 R &= r + 2s', & S &= s - 2r',
 \end{aligned}$$

where (a, b, r, s) are the parameters of the tangent turbines of S and the accent denotes total differentiation with respect to u .

From Theorem 11 and the Corollary to Theorem 11, we obtain

THEOREM 13. *The necessary and sufficient conditions that ∞^1 limaçon series be a set of osculating limaçon series are that they be a set of enveloping limaçon series and their central turbines be a set of tangent turbines in such a way that the element E of the envelope of the limaçon series on any particular limaçon series L is antiparallel (parallel but of opposite orientation) to the element E' of the envelope of the central turbines which is on the central turbine of L .*

THEOREM 14. *The tangent turbines and the central turbines of any general series have in common the envelope of the central turbines.*

The envelope of the central turbines is called the *series of curvature* of the given series. It is given by the equations

$$(32) \quad U = u + \pi, \quad V = v + 2w', \quad W = -2v' + w.$$

THEOREM 15. *There is one and only one general series which contains a given element E_0 and which possesses a given series as series of curvature.*

An osculating limaçon series of a general series S at an element E may be defined as the unique limiting limaçon series of the set of limaçon series such that any limaçon series of this set contains the element E and any other two nearby elements of S .

9. **The osculating circular series of an equiparallel series.** By a process of reasoning similar to that used in the preceding section, we obtain the following results.

THEOREM 16. *The ∞^1 circular series*

$$(33) \quad u = c(t), \quad [v - \alpha(t)]^2 + [w - \beta(t)]^2 = \gamma(t)^2$$

constitute a set of osculating circular series if and only if

$$(34) \quad c' = 0, \quad \alpha'^2 + \beta'^2 = \gamma'^2.$$

COROLLARY. *The series to which the ∞^1 circular series of Theorem 16 are the osculating circular series either consists of one element or is an equiparallel series. It is given by the equations*

$$(35) \quad u = c, \quad v = \alpha - \frac{\alpha'\gamma}{\gamma'}, \quad w = \beta - \frac{\beta'\gamma}{\gamma'}.$$

THEOREM 17. *The osculating circular series of the equiparallel series S of (2) are those whose parameter values are*

$$(36) \quad \alpha = v - \frac{w'(1 + w'^2)}{w''}, \quad \beta = w + \frac{1 + w'^2}{w''}, \quad \gamma = \frac{(1 + w'^2)^{3/2}}{w''},$$

where the accent denotes differentiation with respect to v .

From Theorem 16 and the Corollary to Theorem 16, we obtain

THEOREM 18. *The necessary and sufficient conditions that ∞^1 circular series be an osculating set of circular series are that they all have a common direction and that the circles of the circular series be a set of osculating circles.*

The equiparallel series which has the common direction of the given equiparallel series S and whose point-union is the curve of centers of the osculating circular series of S is called the *series of curvature*. It is given by the equations

$$(37) \quad U = c, \quad V = v - \frac{w'(1 + w'^2)}{w''}, \quad W = w + \frac{1 + w'^2}{w''}.$$

An osculating circular series of an equiparallel series S at the element E may be defined as the unique limiting circular series of the set of circular

series such that any circular series of the set contains the element E and any other two nearby elements of S .

At this point, we note that two series S_1 and S_2 are osculating at a common element E , if and only if they have the same osculating limaçon (or circular) series at E .

10. The curvature and torsion of a general series. The curvature κ at an element E of a general series S is defined by the formula

$$(38) \quad \kappa = (r'^2 + s'^2)^{1/2},$$

where (a, b, r, s) are the parameters of the tangent turbine at E and the acc denotes differentiation with respect to u .

The quantity κ is one-half of the radius of the osculating limaçon series L of S at E ; and also it is one-half of the distance between the centers of the tangent and central turbines of S at E . When the direction is from the center of the tangent turbine to the center of the central turbine, we regard κ as positive. Otherwise, we take it to be negative.

The torsion τ at an element E of a general series S is defined by the formula

$$(39) \quad \tau = \frac{d\bar{u}}{du},$$

where u and \bar{u} are the normal angles of the element E of S and the element \bar{E} , which is the central element of the osculating flat field of S at E , respectively.

It is seen that the torsion τ at an element E of a general series S is the rate of change of the angle of the central element of the osculating flat field per unit radian measure of the angle of the element E .

It is observed that a series is a *whirl series* if and only if its torsion is unity.

From (38) and (39), we find

THEOREM 19. *The curvature $\bar{\kappa}$ of the conjugate series \bar{S} of the general series S is equal to the quotient of the curvature κ and torsion τ of the series S . The torsion $\bar{\tau}$ of the conjugate series \bar{S} of the general series S is the reciprocal of the torsion τ of the series S . That is,*

$$(40) \quad \bar{\kappa} = \frac{\kappa}{\tau}, \quad \bar{\tau} = \frac{1}{\tau}.$$

THEOREM 20. *Two general series which have their curvatures and torsions the same functions of u , the angle between the initial element and any element, are equivalent under the whirl-motion group G_6 .*

Theorem 20 proves that the intrinsic equations of any general series in the geometry of the whirl-motion group G_6 are

$$(41) \quad \kappa = \kappa(u), \quad \tau = \tau(u),$$

where κ is the curvature, τ is the torsion, and u is the angle between the initial element and any element.

It is seen that the necessary and sufficient condition that a general series be co-flat is that its torsion be zero.

Before beginning the proof of Theorem 20, let us consider briefly the feuillet of the plane. Any *feuillet* consists of a lineal element E , a turbine T passing through E , and a flat field F containing both E and T . We recognize three distinct types of feuillet: (1) a *general* feuillet is one where both the turbine T and the flat field F are nonlinear, (2) an *intermediate* feuillet is one where the turbine T is linear and the flat field F is nonlinear, and (3) an *equiparallel* feuillet is one where both the turbine T and the flat field F are linear. The number of general (or intermediate, or equiparallel) feuillet in the plane is ∞^6 (or ∞^4 , or ∞^4).

Under the whirl-motion group G_6 , any two general (or intermediate, or equiparallel) feuillet are equivalent. In particular, under G_6 , any general feuillet can be carried into the general feuillet such that the point and direction of its element E_0 are the origin and the positive direction of the y -axis respectively, its nonlinear turbine T_0 consists of all the lineal elements through the origin (the point-union or the star at the origin), and its nonlinear flat field F_0 is the one whose central element is E_0 . We shall call this the normal feuillet. This result is very important in the proof of our fundamental Theorem 20.

Any feuillet of a general (equiparallel) series S is a general (equiparallel) feuillet which consists of an element E of S , the tangent nonlinear (linear) turbine T to S at E , and the osculating nonlinear (linear) flat field F to S at E . Obviously a general (equiparallel) series S possesses ∞^1 general (equiparallel) feuillet.

We shall now begin the proof of Theorem 20. First, we shall show that there are only two general series S_1 and S_2 with the curvature and torsion given functions of the angle u and with the normal feuillet as initial feuillet. (The angle u is the angle between any element E of S_1 or S_2 and the element E_0 of the normal feuillet.) By (39) and (41), we find

$$(42) \quad \bar{u} = \int_0^u \tau du.$$

By equations (6), (8), (16), (38), and (41), we obtain

$$(43) \quad e^{iu} = -\frac{r' - is'}{a' - ib'}, \quad e^{i\bar{u}} = -\frac{r' + is'}{a' - ib'}, \quad \kappa = (r'^2 + s'^2)^{1/2} = (a'^2 + b'^2)^{1/2},$$

where the accent denotes differentiation with respect to u . Solving these equations for $a' + ib'$ and $r' + is'$, and integrating these results with respect to u , we find

$$(44) \quad a + ib = \mp \int_0^u \kappa e^{i(\bar{u}+u)/2} du, \quad r + is = \pm \int_0^u \kappa e^{i(\bar{u}-u)/2} du,$$

where the upper (or lower) signs are taken simultaneously and where \bar{u} is defined by the equation (42). Since these are the parameters of the tangent turbines of our required series, we find that our two general series S_1 and S_2 are given by

$$(45) \quad v + iw = \mp \left[e^{-iu} \int_0^u \kappa e^{i(\bar{u}+u)/2} du - \int_0^u \kappa e^{i(\bar{u}-u)/2} du \right],$$

where \bar{u} is defined by the equation (42). This establishes our assertion.

Next we shall show that the two series S_1 and S_2 as given by (45) are equivalent under the whirl-motion group G_6 . For the transformation of G_6

$$(46) \quad U = u, \quad V = -v, \quad W = -w,$$

which is the product of a rotation R_π through π radians about the origin by the turn T_π through π radians, converts either one of the two series S_1 and S_2 into the other.

Let S' be any other general series with the curvature and torsion the same functions of the angle u . (The angle u is the angle between any element E and the initial element E' of S' .) Since any two general feuilletts are equivalent under the whirl-motion group G_6 , we can carry the initial general feuillet of S' (determined by the initial element E') into the normal feuillet. Under any such transformation of G_6 , the general series S' is converted into a general series S'' . Since S'' and either one of our original general series S_1 or S_2 possess the same initial feuillet (the normal feuillet) and since their curvatures and torsions are the same functions of the angle u , it follows by what we have proved above that S'' must coincide with either S_1 or S_2 . Hence the three series S' , S_1 , and S_2 are all equivalent to each other under the whirl-motion group G_6 . The proof of Theorem 20 is therefore complete.

11. The curvature of an equiparallel series. The curvature $\kappa = 1/\gamma$ at an element E of an equiparallel series S is defined by the formula

$$(47) \quad \kappa = \frac{1}{\gamma} = \frac{w''}{(1 + w'^2)^{3/2}},$$

where the accent denotes differentiation with respect to v .

The quantity γ is the radius of the osculating circular series C of S at E . When it is the distance from the point of E to the center of the osculating circular series C , we regard $\kappa = 1/\gamma$ as positive. Otherwise, we take it to be negative.

The *torsion* of an equiparallel series is taken to be zero.

THEOREM 21. *Two equiparallel series which have their curvatures the same functions of s , the arc length of the point-union from the initial element to any element, are equivalent under the whirl-motion group G_6 .*

Theorem 21 shows that the intrinsic equations of any equiparallel series in the geometry of the whirl-motion group G_6 are

$$(48) \quad \kappa = \kappa(s), \quad \tau = 0,$$

where κ is the curvature, τ is the torsion, and s is the arc length of the point-union between the initial element and any element.

Theorem 21 is a consequence of the fact that the whirl-motion group G_6 induces the group of rigid motions between the linear flat fields of the plane.

Now we may observe that the curvature of *any* series is the rate of change of the tangent turbine per unit measure of the elements of the series; and the torsion of any series is the rate of change of the osculating flat field per unit measure of the elements of the series.

BIBLIOGRAPHY

1. Kasner, *The group of turns and slides, and the geometry of turbines*, American Journal of Mathematics, vol. 33 (1911), pp. 193-202.
2. Kasner and De Cicco, *The geometry of turbines, flat fields, and differential equations*, American Journal of Mathematics, vol. 59 (1937), pp. 545-563.
3. De Cicco, *The geometry of whirl series*, these Transactions, vol. 43 (1938), pp. 344-358.
4. Kasner and De Cicco, *The geometry of the whirl-motion group G_6 : elementary invariants*, Bulletin of the American Mathematical Society, vol. 44 (1938), pp. 399-403.
5. Kasner and De Cicco, *Quadric fields in the geometry of the whirl-motion group G_6* , American Journal of Mathematics, vol. 61 (1939), pp. 131-142.

BROOKLYN COLLEGE,
BROOKLYN, N. Y.

ON THE REMAINDERS AND CONVERGENCE OF THE SERIES FOR THE PARTITION FUNCTION†

BY
D. H. LEHMER

1. Introduction. The two series under discussion are

$$(1) \quad p(n) = \frac{12^{1/2}}{24n-1} \sum_{k=1}^N A_k^*(n) \left(1 - \frac{k}{\mu}\right) e^{\mu/k} + R_1(n, N),$$

$$(2) \quad p(n) = \frac{12^{1/2}}{24n-1} \sum_{k=1}^N A_k^*(n) \left\{ \left(1 - \frac{k}{\mu}\right) e^{\mu/k} + \left(1 + \frac{k}{\mu}\right) e^{-\mu/k} \right\} + R_2(n, N),$$

due respectively to Hardy and Ramanujan [1]‡ (1917) and to Rademacher [2] (1937). Here we have introduced the abbreviation

$$(3) \quad \mu = \mu(n) = (\pi/6)(24n-1)^{1/2} = O(n^{1/2}).$$

The coefficients A_k^* are real numbers defined by

$$(4) \quad A_k^*(n) = k^{-1/2} A_k(n),$$

where $A_k(n)$ is a complicated sum of $24k$ th roots of unity.§ The remainders $R_1(n, N)$ and $R_2(n, N)$ are defined by (1) and (2) in which $p(n)$ denotes the number of unrestricted partitions of n .

The fact of primary importance about (2) is that

$$(5) \quad \lim_{N \rightarrow \infty} R_2(n, N) = 0;$$

that is to say, the series in (2) as $N \rightarrow \infty$ converges for all n to $p(n)$. Concerning $R_1(n, N)$ Hardy and Ramanujan proved that for every $\alpha > 0$

$$(6) \quad R_1(n, \alpha n^{1/2}) = O(n^{-1/4}).$$

Rademacher [2] gave the following estimate for $R_2(n, N)$ in general:

$$(7) \quad |R_2(n, N)| < \frac{44\pi^2}{225 \cdot 3^{1/2}} N^{-1/2} + \frac{\pi \cdot 2^{1/2}}{75} \frac{N^{1/2}}{(n-1)^{1/2}} \sinh \frac{\pi(2n/3)^{1/2}}{N}$$

and a more complicated estimate for $R_1(n, N)$ from which (6) follows in case $N = \alpha n^{1/2}$. These estimates for the possible errors in (1) and (2) permitted for

† Presented to the Society, February 25, 1939; received by the editors February 24, 1939.

‡ The numbers in square brackets refer to the papers listed in the bibliography at the end of this paper.

§ For a complete definition of the A 's see either [1, p. 85], [2, p. 242], or [3, pp. 271-273].

the first time the use of either (1) or (2) with absolute assurance. Using the estimate

$$(8) \quad |A_k(n)| < 2k^{5/6}$$

instead of the trivial

$$(9) \quad |A_k(n)| < k$$

previously employed, the writer obtained [4, 5]

$$(10) \quad |R_2(n, N)| < \frac{\pi^2}{3^{1/2}} N^{-2/3} \left\{ \left(\frac{N}{\mu} \right)^3 \sinh \frac{\mu}{N} + \frac{1}{6} - \left(\frac{N}{\mu} \right)^2 \right\},$$

$$(11) \quad |R_1(n, N)| < \frac{\pi^2 N^{7/3}}{3^{1/2} \mu^3} \left\{ \sinh \frac{\mu}{N} + \frac{1}{6} \left(\frac{\mu}{N} \right)^3 + \left(1 + \frac{N}{\mu} \right) \left(\frac{1}{7} + \frac{1}{3} \mu^{1/3} N^{-5/3} \right) \right\}.$$

If in (10) and (11) we substitute $N = \alpha n^{1/2}$, we find that in either case

$$(12) \quad R_i(n, \alpha n^{1/2}) = O(n^{-1/3}) \quad (i = 1, 2).$$

In §2 we show by a simple asymptotic argument that

$$(13) \quad R_i(n, \alpha n^{1/2}) = O(n^{-1/2} \log n) \quad (i = 1, 2),$$

a result, which in a sense, is the best possible. In §3 by a more precise treatment we obtain formulas similar to (10) and (11) but of which (13) rather than (12) is a special case.

Hardy and Ramanujan [1, p. 107] raised the question of the boundedness of $A_k(n)$ in discussing the possible convergence of (1) as $N \rightarrow \infty$. In proving the divergence of (1) the writer [6] employed a sequence of A 's which, if they tended to zero, did not do so rapidly enough to render (1) convergent. Although this showed, in other words, that $R_1(n, N)$ tends to zero for no value of n , it did not remove the possibility of $R_1(n, N)$ ultimately oscillating between fixed limits. Incidentally to this discussion it was shown that $A_k(0)$ and $A_k(-1)$ are unbounded. Later [4, Theorem 11], it was proved that $A_k(n)$ is an unbounded function of k for infinitely many values of n . In §4 we show that this is true for every value of $n > 0$ or $n < 0$, proving in fact that for all n $A_k(n) = \Omega(k^{1/2})$. (The interest in $A_k(n)$ is not confined to positive values of n [3, p. 83; 7, p. 466].) From this result it follows that $R_1(n, N)$ does not oscillate between fixed limits, the terms of the series in (1) being unbounded. It follows also that the k th term of (2) is greater in absolute value than $1/k^2$ for an infinity of k 's despite the apparent rapidity of its convergence.

The writer wishes to acknowledge several helpful suggestions of Dr. H. Heilbronn especially in connection with Lemma 4.

2. **Proof of (13).** It is convenient to begin with

LEMMA 1. *If α is a positive constant, then for $s < 1$,*

$$(14) \quad \sum_{k \leq \alpha n^{1/2}} A_k^*(n) k^{-s} = O(n^{(1-s)/2} \log n),$$

and if $s > 1$, then

$$(15) \quad \sum_{k > \alpha n^{1/2}} A_k^*(n) k^{-s} = O(n^{(1-s)/2} \log n).$$

Proof. By Theorem 8 of [4], $A_k^*(n)$ in absolute value does not exceed $2^{\omega(k)}$, the number of odd quadratfrei divisors of k , and hence does not surpass $\tau(k)$, the number of divisors of k . If therefore we denote, as usual, by $T(k)$ the sum function

$$(16) \quad T(k) = \sum_{v=1}^k \tau(v) = O(k \log k),$$

then we have

$$\begin{aligned} \sum_{a < k \leq b} |A_k^*(n)| k^{-s} &= O\left(\sum_{a < k \leq b} \tau(k) k^{-s}\right) = O\left(\sum_{a < k \leq b} T(k) \{k^{-s} - (k+1)^{-s}\}\right) \\ &= O\left(\sum_{a < k \leq b} k(\log k) k^{-s-1}\right) = O\left(\sum_{a < k \leq b} k^{-s} \log k\right) \\ &= O\left(\int_a^b x^{-s} \log x dx\right) = O(b^{1-s} \log b - a^{1-s} \log a). \end{aligned}$$

To prove equation (14) set $a=1$ and $b=\alpha n^{1/2}$. To prove (15) set $a=\alpha n^{1/2}$, $b=\infty$.

THEOREM 1. $R_2(n, \alpha n^{1/2}) = O(n^{-1/2} \log n)$.

Proof. If we expand the exponentials in (2) and collect the terms, we have

$$(17) \quad R_2(n, N) = \frac{4 \cdot 12^{1/2}}{24n-1} \sum_{k=N+1}^{\infty} A_k^*(n) \sum_{j=1}^{\infty} \frac{j(\mu/k)^{2j}}{(2j+1)!}.$$

Hence in view of (3)

$$R_2(n, \alpha n^{1/2}) = O\left(\frac{1}{n} \sum_{j=1}^{\infty} \frac{j n^j}{(2j+1)!} \sum_{k > \alpha n^{1/2}} |A_k^*(n)| k^{-2j}\right).$$

Applying (15) with $s=2j$, we have

$$R_2(n, \alpha n^{1/2}) = O\left(\frac{1}{n} \sum_{j=1}^{\infty} \frac{j n^j}{(2j+1)!} n^{(1-2j)/2} \log n\right) = O(n^{-1/2} \log n).$$

It remains to prove

THEOREM 2. $R_1(n, \alpha n^{1/2}) = O(n^{-1/2} \log n)$.

Proof. Let $D(n, N)$ represent the difference between the sum of the first N terms of (1) and the first N terms of (2). Then in view of Theorem 1 it suffices to show that $D(n, \alpha n^{1/2}) = O(n^{-1/2} \log n)$. Now

$$D(n, \alpha n^{1/2}) = \frac{12^{1/2}}{24n-1} \sum_{k \leq \alpha n^{1/2}} A_k^*(n) \left(1 + \frac{k}{\mu}\right) e^{-\mu/k}.$$

Since $e^{-\mu/k} < k/\mu$, we have by (3)

$$D(n, \alpha n^{1/2}) = O\left(\frac{1}{n} \left\{ \sum_{k \leq \alpha n^{1/2}} |A_k^*(n)| \frac{k}{n^{1/2}} + \sum_{k \leq \alpha n^{1/2}} |A_k^*(n)| \frac{k^2}{n} \right\}\right).$$

Applying (14) with $s = -1$ and $s = -2$, we have

$$D(n, \alpha n^{1/2}) = O(n^{-3/2} n \log n + n^{-2} n^{3/2} \log n) = O(n^{-1/2} \log n).$$

This completes the proof of (13).

3. Estimates of the general remainders $R_i(n, N)$. In what follows we shall use the function $T(n)$ as before but will require something more explicit than (16). Hence we start with

LEMMA 2. For all positive integers k

$$(18) \quad T(k) > k \log k,$$

while for $k > 12$

$$(19) \quad T(k) \leq k(\log k + b_{16})$$

where $b_{16} = T(16)/16 - \log 16 = .3524113 \dots$

Proof. We shall need the following inequalities:

$$(20) \quad \log x - \delta/(x - \delta) < \log(x - \delta) < \log x - \delta/x, \quad 0 < \delta < 1 < x,$$

$$(21) \quad \log m + C < H(m) < \log m + C + 1/(2m)$$

where $H(m) = 1 + 1/2 + 1/3 + \dots + 1/m$ and $C = .577215 \dots$ is Euler's constant. The inequalities (20) follow readily from $e^{\delta/(x-\delta)} > 1/(1-\delta/x) > e^{\delta/x}$, which are seen at once to be true on expanding the functions involved, while (21) follows from the familiar asymptotic expansion

$$H(m) = \log m + C + 1/(2m) - 1/(12m^2) - \dots$$

If in the well known relation

$$(22) \quad T(k) = 2 \sum_{x \leq k^{1/2}} \left[\frac{k}{x} \right] - [k^{1/2}]^2$$

we remove the greatest integer signs in the sum, we obtain by (21)

$$(23) \quad T(k) \leq 2kH([k^{1/2}]) - [k^{1/2}]^2 < 2k \log [k^{1/2}] + 2Ck + k/[k^{1/2}] - [k^{1/2}]^2.$$

Writing

$$(24) \quad [k^{1/2}] = k^{1/2} - \delta$$

and applying (20) we have at once from (23)

$$\begin{aligned} T(k) &< k \log k + (2C - 1)k + k/(k^{1/2} - \delta) - \delta^2 \\ &< k(\log k + 2C - 1 + 1/(k^{1/2} - 1)). \end{aligned}$$

Now if $k \geq 37$, then

$$2C - 1 + 1/(k^{1/2} - 1) < .15444 + 1/(37^{1/2} - 1) < .35122 < b_{16}.$$

Hence (19) is true for $k > 36$. That it is true for $12 < k \leq 36$ is shown by the following table of $T(k)$ and

$$(25) \quad b_k = T(k)/k - \log k$$

which will be of use later.

TABLE I

k	$T(k)$	b_k	k	$T(k)$	b_k	k	$T(k)$	b_k	k	$T(k)$	b_k
1	1	1.000	11	29	.238	21	70	.289	31	113	.211
2	3	.807	12	35	.432	22	74	.273	32	119	.253
3	5	.568	13	37	.238	23	76	.169	33	123	.231
4	8	.614	14	41	.288	24	84	.322	34	127	.209
5	10	.391	15	45	.292	25	87	.261	35	131	.188
6	14	.541	16	50	.352	26	91	.242	36	140	.305
7	16	.340	17	52	.225	27	95	.186	37	142	.227
8	20	.421	18	58	.332	28	101	.275	38	146	.168
9	23	.358	19	60	.213	29	103	.184	39	150	.183
10	27	.397	20	66	.304	30	111	.299	40	158	.261

It is seen that (19) is true also for $k=7$ and 11 and that the equal sign holds only when $k=16$. Any number of inequalities similar to (19), such as

$$T(k) < k(\log k + b_{18}) < k(\log k + .33185047), \quad k > 16,$$

may be established in the same way.

To prove (18) we use the inequality

$$[k/x] \geq (k - x + 1)/x,$$

so that (22) gives

$$\begin{aligned} T(k) &\geq 2 \sum_{x \leq [k^{1/2}]} \frac{k+1}{x} - 2[k^{1/2}] - [k^{1/2}]^2 \\ &= 2(k+1)H([k^{1/2}]) - 2[k^{1/2}] - [k^{1/2}]^2. \end{aligned}$$

By (20), (21) and (24) we have therefore

$$\begin{aligned} T(k) &> (k+1)(\log k - 2\delta/(k^{1/2} - \delta) + 2C) - (1 + [k^{1/2}])^2 + 1 \\ &> (k+1) \log k + (2C-1)k + 2C - 2k^{1/2} - 2(k^{1/2} + 1)/(k^{1/2} - 1) \\ &= k \log k + (2C-1)k + \log k - 2((k+1)/(k^{1/2} - 1) - C). \end{aligned}$$

The function $(2C-1)k + \log k - 2((k+1)/(k^{1/2}-1) - C) > 0$ for $k > 117$. Hence (18) is true, if true for $k \leq 117$, and this is readily verified. In fact in view of (25) it is seen that (18) is equivalent to $b_k > 0$, an inequality which holds for $k \leq 117$, the smallest value of b_k being $b_{99} = .14280154$. Of course b_k tends to $2C-1 = .1544 \dots$ as $k \rightarrow \infty$.

LEMMA 3. Let $N > 12$, $s > 1$, and let n be any integer. Then

$$\begin{aligned} (26) \quad \sum_{k=N}^{\infty} |A_k^*(n)| k^{-s} &< -T(N-1)N^{-s} \\ &\quad + s(s-1)^{-1}N^{1-s} \{ \log N + (s-1)^{-1} + .3524113 \}. \end{aligned}$$

Proof. From the fact that

$$(27) \quad |A_k^*(n)| \leq \tau(k)$$

it follows that

$$\begin{aligned} \sum_{k=N}^{\infty} |A_k^*(n)| k^{-s} &\leq \sum_{k=N}^{\infty} \tau(k) k^{-s} = \sum_{k=N}^{\infty} T(k) \{ k^{-s} - (k+1)^{-s} \} - T(N-1)N^{-s} \\ &< -T(N-1)N^{-s} - \int_N^{\infty} T(x) d(x^{-s}) \\ &= -T(N-1)N^{-s} + s \int_N^{\infty} T(x) x^{-s-1} dx. \end{aligned}$$

Here we have defined $T(x)$ as $T([x])$. Applying Lemma 1 with $k=N > 12$ and integrating, we obtain the lemma at once.

THEOREM 3. If $N > 12$, then

$$\begin{aligned} R_2(n, N-1) &< (48^{1/2}\pi^2/9N) \{ w_1(\mu/N)(\log N + .3524113) + w_2(\mu/N) \\ &\quad - (T(N-1)/2N)w_3(\mu/N) \} \end{aligned}$$

where the functions $w_i(x)$ are defined by

$$(28) \quad \begin{aligned} w_1(x) &= \sum_{j=1}^{\infty} \frac{j^2 x^{2j-2}}{(2j-1)(2j+1)!}, & w_2(x) &= \sum_{j=1}^{\infty} \frac{j^2 x^{2j-2}}{(2j-1)^2(2j+1)!}, \\ w_3(x) &= \sum_{j=1}^{\infty} \frac{j x^{2j-2}}{(2j+1)!}. \end{aligned}$$

Proof. By (17) we have

$$R_2(n, N-1) = \frac{4 \cdot 12^{1/2}}{24n-1} \sum_{j=1}^{\infty} \frac{j \mu^{2j}}{(2j+1)!} \sum_{k=N}^{\infty} A_k^*(n) k^{-2j}.$$

Taking absolute values and applying Lemma 3 with $s=2j$, we find that

$$\begin{aligned} |R_2(n, N-1)| &< \frac{4 \cdot 12^{1/2} \mu^2}{(24n-1)N} \left\{ -\frac{T(N-1)}{N} \sum_{j=1}^{\infty} \frac{j(\mu/N)^{2j-2}}{(2j+1)!} \right. \\ &\quad \left. + 2(\log N + b_{16}) \sum_{j=1}^{\infty} \frac{j^2(\mu/N)^{2j-2}}{(2j-1)(2j+1)!} + 2 \sum_{j=1}^{\infty} \frac{j^2(\mu/N)^{2j-2}}{(2j-1)^2(2j+1)!} \right\}. \end{aligned}$$

Making use of (3) and (28), we obtain the theorem at once.

THEOREM 4. *If N is any positive integer,*

$$\begin{aligned} |R_1(n, N)| &< (72 \cdot 3^{1/2} / \pi^2) (N/\mu)^2 e^{-\mu/N} \left\{ (T(N)/N^2) (1 + (N+1)/\mu) \right. \\ &\quad \left. - (1/2\mu)(\log N - 1/2) \right\} + |R_2(n, N)|. \end{aligned}$$

Proof. As before let $D(n, N)$ denote the difference between the sums of the first N terms of (1) and (2) so that

$$D(n, N) = \frac{12^{1/2}}{24n-1} \sum_{k=1}^N A_k^*(n) \left(1 + \frac{k}{\mu} \right) e^{-\mu/k}.$$

Then

$$(29) \quad |R_1(n, N)| < |R_2(n, N)| + |D(n, N)|.$$

Now

$$(30) \quad |D(n, N)| < \frac{12^{1/2} e^{-\mu/N}}{24n-1} \left\{ \sum_{k=1}^N |A_k^*(n)| + \frac{1}{\mu} \sum_{k=1}^N k |A_k^*(n)| \right\}$$

but from (27)

$$\sum_{k=1}^N |A_k^*(n)| \leq T(N)$$

while

$$(31) \quad \sum_{k=1}^N k |A_k^*(n)| \leq \sum_{k=1}^N k \tau(k) = (N+1)T(N) - \sum_{k=1}^N T(k).$$

By (18) we have

$$\sum_{k=1}^N T(k) > \sum_{k=1}^N k \log k > \int_1^N x \log x dx > \frac{1}{2} N^2 \log N - \frac{N^2}{4},$$

so that by (30) and (31)

$$|D(n, N)| < \frac{12^{1/2} e^{-\mu/N}}{24n-1} \left\{ T(N) + \frac{N+1}{\mu} T(N) - \frac{N^2}{4\mu} (2 \log N - 1) \right\}.$$

The theorem now follows from (3) and (29).

Of the three functions (28) only $w_3(x)$ is elementary; in fact

$$w_3(x) = \frac{1}{2x^2} (x \cosh x - \sinh x),$$

the other two depending on higher transcendents. For our purposes it is best to use their series not only as definitions but also as effective means of evaluating these functions. A short table of $w_1(x)$, $w_2(x)$ and $w_3(x)$ is given below.

TABLE II

x	$w_1(x)$	$w_2(x)$	$w_3(x)$
1	.1781	.1704	.1839
2	.2172	.1827	.2436
3	.3007	.2065	.3738
4	.4668	.2485	.6402
5	.7992	.3215	1.1874
6	1.4794	.4535	2.3347
7	2.9089	.6873	4.7958
8	5.9877	1.1327	10.1901
9	12.7652	1.9991	22.2307
10	27.9660	3.7336	49.5596

In actual practice we are concerned with $n > 600$, since tables of $p(n)$ now extend to $p(600)$. Unless we carry the calculation to a considerable number of places to the right of the decimal point and at the same time employ quite a large number of terms, we cannot distinguish between the terms of (1) and (2). Hence in practice we may use (1) and apply Theorem 3 to estimate the remainder. We give three examples of the application of above estimates

TABLE III

	By (7)	By (10)	Theorem 3	Actual value	μ
$ R_2(721, 21) $.378	.341	.231	.00041	68.875
$ R_3(2052, 18) $	2.028	1.099	.815	.0408	116.20
$ R_4(14031, 63) $.387	.245	.150	.00016	303.84

in widely different cases. Linear interpolation may be used for $w_1(x)$ and $w_2(x)$, since it will give values in excess of the actual values of these functions.

Values of $T(n)$ can be taken from Table I for $n \leq 40$ and can be quickly found from (22) if $n > 40$. For rough calculation we may use the inequality

$$T(N-1)/2N > \log N^{1/2}.$$

If we replace N by $\alpha n^{1/2}$ so that

$$\mu/N = O(1), \quad T(N)/N = O(n^{-1/2} \log n)$$

in Theorems 3 and 4, it is seen that (13) is a special case of these theorems.

If instead of (27) we were to use

$$(32) \quad |A_k^*(n)| \leq 2^{\omega(k)},$$

then for very large values of N it would be possible to obtain smaller estimates for $|R_i(n, N)|$ by a simple modification of the above argument. In fact we would then be concerned with the function

$$(33) \quad \psi(k) = \sum_{r=1}^k 2^{\omega(r)} = \frac{3}{\pi^2} k \log k + O(k),$$

so that theoretically one could reduce the estimate for $|R_i(n, N)|$ by a factor of nearly $3/\pi^2$. This of course would not alter (13). If one is to use some inequality for $|A_k^*(n)|$ of the same type as (27) or (32) in which the right side is independent of n , then it is impossible to obtain an essentially better inequality than (32). In this sense (13) cannot be improved upon. In practice a small bound for the constant implied in (33) is not easily obtained, nor indeed can one achieve the factor $\dagger 3/\pi^2$. In the end one obtains theorems similar to Theorems 3 and 4 which are superior only for larger values of N than one would naturally encounter in actual calculations.

4. Proof of unboundedness of $A_k(n)$. In proving that $A_k(n)$ is unbounded it is necessary to consider separately the cases in which n is and is not the negative of a pentagonal number. In the first case the proof is quite simple. In the second case we make use of a lemma depending on the prime number theorem.

THEOREM 5. *If $-n$ is a pentagonal number, there exist infinitely many primes p such that $|A_p(n)| > (3p)^{1/2}$. This is not true for a larger number than 3.*

Proof. By Theorem 5 of [4]

$$(34) \quad |A_p(n)| = 2p^{1/2} |\cos(4\pi m/p)|$$

where the integer m satisfies the congruence

$$(35) \quad (24m)^2 \equiv 1 - 24n \pmod{p}.$$

\dagger One can prove for instance that $\psi(k) < .6534k \log k + 3.387k$.

If $-n$ is a pentagonal number so that $n = -(3u^2 \pm u)/2$ or rather $1 - 24n = (6u \pm 1)^2$, the congruence (35) has for every prime $p > 3$ not dividing $1 - 24n$ the integral solution $m = (6u \pm 1)(p^2 - 1)/24$. Hence by (34)

$$\begin{aligned} |A_p(n)| &= 2p^{1/2} \left| \cos \left\{ (6u \pm 1)(p - 1/p)\pi/6 \right\} \right| \\ &= 2p^{1/2} \left| \cos \left\{ (p - [6u \pm 1]/p)\pi/6 \right\} \right|. \end{aligned}$$

As $p \rightarrow \infty$ through prime values this tends steadily to $2p^{1/2} \cos \pi/6 = (3p)^{1/2}$ and approaches it from above or below according as $p \rightarrow \infty$ through values of the form $6k+1$ or $6k-1$. Thus the assertions of the theorem are proved.

LEMMA 4. Let a and b be coprime integers such that $-ab$ is a non-square and a is even, and let t_M denote the least positive solution (if it exists) of the congruence

$$at^2 + b \equiv 0 \pmod{M}.$$

Finally let γ be a constant greater than $1/2$; then the inequality

$$t_p < \gamma p / \log p$$

holds for infinitely many primes p .

Proof. Let \sum' denote a summation over those primes greater than 2 of which $-ab$ is a quadratic residue. We recall that asymptotically half the primes are of this sort. Let x be a large integer. Then identically

$$(36) \quad \sum_{k=1}^x \log |ak^2 + b| = \sum' (\log p) \left\{ \left[\frac{x + p - t_p}{p} \right] + \left[\frac{x + t_p}{p} \right] + \sum_{v=2}^{\infty} \left(\left[\frac{x + p^v - t_{p^v}}{p^v} \right] + \left[\frac{x + t_{p^v}}{p^v} \right] \right) \right\}.$$

Now

$$\sum_{k=1}^x \log |ak^2 + b| = 2x \log x + O(x).$$

The right member of (36) may be written

$$\begin{aligned} \sum'_{p < 2x} \left(\frac{2x}{p} + 1 \right) \log p + O \left(\sum'_{p < 2x} \log p \right) &+ \sum'_{p^v < ax^2 + b, v > 1} \left(\frac{2x}{p^v} + 1 \right) \log p \\ &+ \sum'_{p > 2x, t_p < x} \log p + O(x) = 2x \sum'_{p < 2x} \log p / p + \sum'_{p > 2x, t_p < x} \log p + O(x) \\ &= x \log x + \sum_{p > 2x, t_p < x} \log p + O(x). \end{aligned}$$

Now suppose that the lemma is false so that $t_p \geq \gamma p / \log p$ for all suffi-

ciently large p . Then for a sufficiently large p the inequality $t_p < x$ implies $x > \gamma p / \log p$. That is

$$p < (1/\gamma)x \log p = (1/\gamma)x \log x + O(x \log \log x).$$

Hence

$$\begin{aligned} \sum'_{p > 2x, t_p < x} \log p &= \sum'_{p > 2x, p < x(\log x)/\gamma} \log p + O(x \log \log x) \\ &= (x/2\gamma) \log x + O(x \log \log x). \end{aligned}$$

Therefore (36) may be written for all sufficiently large x

$$2x \log x = x \log x + (1/2\gamma)x \log x + O(x \log \log x)$$

or

$$(2\gamma - 1)x \log x = O(x \log \log x).$$

But as $\gamma > 1/2$, this is a contradiction. Hence the lemma is true.

The proof of this lemma was suggested to the writer by Dr. H. Heilbronn. As a matter of fact the hypotheses of the lemma are unnecessarily restrictive but are sufficient to meet our immediate needs. By only a slight complication of (36) the same proof applies to any irreducible quadratic.

THEOREM 6. *If $-n$ is not a pentagonal number, there exist for every $\epsilon > 0$ infinitely many primes p such that*

$$(37) \quad |A_p(n)| > (2 - \epsilon)p^{1/2}.$$

Proof. In Lemma 4 choose $a = 24^2$ and $b = 24n - 1$. This is permissible since a is even and prime to b , and $-ab = 24^2(1 - 24n)$ is a non-square because $-n$ is not a pentagonal number. Then by Lemma 4 there exist infinitely many primes p for which the congruence (35) has a solution m such that $0 < m < \eta p$ where η is a positive constant less than $1/8$ to be determined presently. Then for every such p , by (34),

$$|A_p(n)| > 2p^{1/2} \cos 4\pi\eta.$$

To obtain (37) one has only to choose η so small that $\cos 4\pi\eta$ differs from unity by less than $\epsilon/2$.

THEOREM 7. *For every positive n the k th term of the Hardy-Ramanujan series (1) is in absolute value greater than*

$$13k(24n - 1)^{-3/2}$$

for infinitely many values of k .

Proof. By Theorem 6 for every $\epsilon > 0$ there exist infinitely many primes p

for which the p th term of (1) is greater in absolute value than

$$\frac{12^{1/2}}{24n-1} \left| 1 - \frac{p}{\mu} \right| e^{\mu/p}(2-\epsilon) > \frac{6 \cdot 12^{1/2} p(2-\epsilon)}{\pi(24n-1)^{3/2}} - \frac{2\epsilon 12^{1/2}}{24n-1}$$

provided $p > \mu$. There exists a positive ϵ so small that $6 \cdot 12^{1/2}(2-\epsilon)/\pi > 13$. For such an ϵ and for all sufficiently large primes p associated with this ϵ , the p th term of (1) is greater than $13p(24n-1)^{-3/2}$.

THEOREM 8. *For every positive n the k th term of the Rademacher series (2) is in absolute value greater than $(43/34)k^{-2}$ for infinitely many k 's.*

Proof. Since

$$(1-x^{-1})e^x + (1+x^{-1})e^{-x} = 2(x^2/3 + x^4/30 + \dots) > (2/3)x^2,$$

the p th term of (2) in view of (3), (4) and (37) is, in absolute value, greater than

$$\frac{12^{1/2}}{24n-1} \cdot \frac{2}{3} \frac{\mu^2}{p^2} (2-\epsilon) = \frac{2\pi^2}{9 \cdot 3^{1/2}} p^{-2} - \epsilon' p^{-2} > \frac{43}{34} p^{-2}$$

for a suitably chosen ϵ .

BIBLIOGRAPHY

1. G. H. Hardy and S. Ramanujan, *Asymptotic formulas in combinatory analysis*, Proceedings of the London Mathematical Society, (2), vol. 17 (1918), pp. 75-115; also S. Ramanujan, *Collected Papers*, pp. 276-309.
2. H. Rademacher, *On the partition function $p(n)$* , Proceedings of the London Mathematical Society, (2), vol. 43 (1937), pp. 241-254.
3. ———, *A convergent series for the partition function $p(n)$* , Proceedings of the National Academy of Sciences, vol. 23 (1937), pp. 78-84.
4. D. H. Lehmer, *On the series for the partition function*, these Transactions, vol. 43 (1938), pp. 271-295.
5. ———, *An application of the Schläfli's modular equation to a conjecture of Ramanujan*, Bulletin of the American Mathematical Society, vol. 44 (1938), pp. 84-90.
6. ———, *On the Hardy-Ramanujan series for the partition function*, Journal of the London Mathematical Society, vol. 12 (1937), pp. 171-176.
7. H. Petersson, *Linear relationen zwischen Poincareschen Reihen*, Abhandlung des mathematischen Seminars der Hamburgischen Universität, vol. 12 (1938), pp. 415-472.

LEHIGH UNIVERSITY,
BETHLEHEM, PA.

CONTRIBUTIONS TO THE TRANSFORMATION THEORY OF DYNAMICS*

BY

DANIEL C. LEWIS, JR.†

The applications of transformation theory to dynamics are familiar through the writings of Poincaré, Levi-Civita, Hadamard, Birkhoff and others. The last named mathematician has given an extensive treatment of the so-called conservative transformations which are of particular use in the study of dynamical systems of two degrees of freedom.‡ Our present purpose is to initiate a similar treatment for those transformations in spaces of higher dimensions which are particularly important from the dynamical point of view and may be regarded as appropriate generalizations of the conservative surface transformations. These are the so-called Pfaffian transformations as defined in §1. In this paper we restrict attention to properties which are essentially characteristic of the Pfaffian transformations and not properties which the Pfaffian transformations have in common with other transformations.§ The most important results of the present paper were discovered independently by G. D. Birkhoff.|| The proofs are published now for the first time.

1. Definition of a Pfaffian transformation and some elementary theorems. Consider a region R of n dimensional space in which are defined n analytic functions $X_1(x), \dots, X_n(x)$ of the n variables x_1, \dots, x_n . For the sake of brevity we write

$$a_{ij} = \partial X_i / \partial x_j - \partial X_j / \partial x_i, \quad i, j = 1, 2, \dots, n,$$

and we assume that the skew symmetric determinant $|a_{ij}|$ is of rank $2k$ at every point of R , where k is a positive integer not greater than $n/2$.

Consider also an analytic transformation T , and let us denote by $\bar{x}_1, \dots, \bar{x}_n$ [or more briefly by (\bar{x})] the point into which the point (x) is carried by T . It is assumed that T is defined when (x) is in R .

* Presented to the Society, April 19, 1935 and September 8, 1939; received by the editors March 31, 1939.

† These results were obtained for the most part while the writer was a National Research Fellow.

‡ G. D. Birkhoff [1]. The numbers in brackets refer to the list of references at end of paper.

§ For a treatment of some of these latter properties cf. Lewis [1, 2].

|| Birkhoff [2, p. 144]. My own results were presented to the American Mathematical Society April 19, 1935 (cf. Lewis [3]) several months before Professor Birkhoff's paper was off the press, but more than a year after his results were obtained. I first learned of his results at the meeting after I had presented my own.

DEFINITION. The transformation T is said to be Pfaffian with respect to the linear differential form

$$(1.1) \quad \sum_{i=1}^n X_i(x) dx_i,$$

if $\sum_{i=1}^n [X_i(\bar{x}) d\bar{x}_i - X_i(x) dx_i]$, thought of as a differential form in the n independent variables x_1, \dots, x_n , is an exact differential, at least whenever (\bar{x}) as well as (x) belong to R .

If two linear differential forms differ from each other by an exact differential, then a transformation which is Pfaffian with respect to one of them is also Pfaffian with respect to the other. Linear differential forms differing from each other by exact differentials shall therefore be said to be equivalent to each other.* It should be further noted that equivalent differential forms give rise to identical matrices (a_{ij}) .

If

$$\sum_{i=1}^n [X_i(\bar{x}) d\bar{x}_i - X_i(x) dx_i] = \sum_{i=1}^n \left[\sum_{j=1}^n X_j(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_i} - X_i(x) \right] dx_i$$

is to be an exact differential, it is necessary and sufficient that

$$\frac{\partial}{\partial x_h} \left\{ \sum_{j=1}^n X_j(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_i} - X_i(x) \right\} = \frac{\partial}{\partial x_i} \left\{ \sum_{j=1}^n X_j(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_h} - X_h(x) \right\},$$

$i, h = 1, \dots, n.$

Hence we obtain the result that a necessary and sufficient condition that T be Pfaffian with respect to (1.1) is that

$$(1.2) \quad \sum_{j=1}^n a_{ji}(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_i} \frac{\partial \bar{x}_i}{\partial x_h} = a_{ih}(x), \quad i, h = 1, \dots, n.$$

Another necessary and sufficient condition that T be Pfaffian with respect to (1.1) is that $\int \sum_{i=1}^n X_i dx_i$ be a relatively invariant integral of T . That is, if C is an arbitrary closed regular curve in R which is carried by T into a curve \bar{C} also in R , then the above line integral extended over C is equal to the integral extended over \bar{C} in the corresponding sense.†

That there exist infinitely many Pfaffian transformations corresponding to an arbitrary linear differential form (1.1) may be shown in the following way:

* This is not the usual definition of equivalence for two differential forms. Cf. Weber [1, p. 128].

† There are in general numerous other integral invariants which can be written in compact form in the notation of Grassman's "exterior calculus." Cf. Goursat [1, pp. 211-212, 229-235]. They are derived from the relatively invariant line integral exactly as for differential equations.

Consider a system of analytic differential equations of the type

$$(1.3) \quad \frac{dx_i}{dt} = F_i(x), \quad i = 1, \dots, n.$$

The right-hand members, for simplicity, may be assumed not to involve t explicitly. Let $x_i = f_i(t, \bar{x})$ be the solution, which for $t=0$ takes on the initial values $x_i = \bar{x}_i$. These equations, for every fixed value of t , may be regarded as defining a transformation from the point (\bar{x}) to the point (x) . It is known that a necessary and sufficient condition that (1.3) admit $\int \sum_{i=1}^n X_i dx_i$ as a relatively invariant line integral is that $\sum_{i,j} a_{ij} F_j dx_i$ be an exact differential.* It may be easily proved that the F 's can always be chosen in an infinite number of ways so that this condition is fulfilled. The rest of the proof is left to the reader.

2. The partial reduction of the form (1.1). It is known from the theory of linear differential forms that, if $2k < n$, we may introduce a change of variables, valid in the neighborhood of an assigned point, in such a way that the form (1.1), or one equivalent to it, may be written as a form in a smaller number of variables.† By repeated application of this result, we may without loss of generality assume that (1.1) is a form in just $2k$ variables, at least if we restrict attention to the vicinity of an invariant point of T .

Let us now suppose that the variables x_1, \dots, x_n are such that the differential form (1.1) involves only x_1, \dots, x_{2k} . In other words $X_\alpha = 0$, ($\alpha = 2k+1, \dots, n$), and $\partial X_i / \partial x_\alpha = 0$, ($i = 1, \dots, n$; $\alpha = 2k+1, \dots, n$), and hence $a_{i\alpha} = -a_{\alpha i} = 0$. It follows that the transformation T (assumed to be non-singular) may be written in the form

$$(2.1) \quad \begin{aligned} \bar{x}_l &= \bar{x}_l(x_1, \dots, x_{2k}), & l &= 1, \dots, 2k, \\ \bar{x}_\alpha &= \bar{x}_\alpha(x_1, \dots, x_{2k}, x_{2k+1}, \dots, x_n), & \alpha &= 2k+1, \dots, n. \end{aligned}$$

In other words the first $2k$ of the equations defining T are independent of x_{2k+1}, \dots, x_n . The variables x_1, \dots, x_{2k} are thus said to form a "separated system."

To prove the italicized statement, we note from (1.2) that under the present hypotheses

$$(2.2) \quad \sum_{j,l=1}^{2k} a_{jl}(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_i} \frac{\partial \bar{x}_l}{\partial x_h} = a_{ih}(x), \quad i, h = 1, 2, \dots, 2k,$$

$$(2.3) \quad \sum_{j,l=1}^{2k} a_{jl}(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_i} \frac{\partial \bar{x}_l}{\partial x_\alpha} = 0, \quad \alpha = 2k+1, \dots, n.$$

* Goursat [1, p. 219].

† Weber [1, pp. 216-217].

Here the $4k^2$ scalar equations (2.2) may be thought of as a single matrix equation, the matrix on the right being the $2k$ -rowed skew symmetric matrix (a_{ih}) whose rank by hypothesis is $2k$. Hence the determinant, the element in whose i th row and l th column is $\sum_{j=1}^{2k} a_{jl}(\bar{x}) \partial \bar{x}_j / \partial x_i$, cannot be zero. It follows from (2.3) that $\partial \bar{x}_l / \partial x_a = 0$, ($l = 1, \dots, 2k$).

If $2k = n$, we shall say that the Pfaffian transformation is *nondegenerate*. The significance of the theorem just proved is that, in studying the characteristic properties of Pfaffian transformations, we may for the most part confine attention to the nondegenerate case.

Apart from the fact that the form (1.1) has not yet been considered in its canonical form, the result obtained above is closely analogous to a theorem of Lie and Koenig on differential equations.*

3. Invariant manifolds. Let T be a nondegenerate Pfaffian transformation, represented by the equations $\bar{x}_i = \bar{x}_i(x)$, ($i = 1, \dots, n = 2k$). Let it be assumed that T admits an m -dimensional analytic invariant manifold M , given by the equations $x_i = f_i(u_1, \dots, u_m)$. This means that there exists a nonsingular transformation S in m -dimensional space of the type $\bar{u}_l = \bar{u}_l(u)$, ($l = 1, 2, \dots, m$), such that

$$f_i(\bar{u}_1, \dots, \bar{u}_m) \equiv \bar{x}_i[f(u_1, \dots, u_m)], \quad i = 1, \dots, 2k,$$

the identity being with respect to u_1, \dots, u_m . It is obvious that S is also a Pfaffian transformation, with respect to the linear differential form

$$\sum_{l=1}^m \left(\sum_{i=1}^{2k} X_i[f(u_1, \dots, u_m)] \frac{\partial f_i}{\partial u_l} \right) du_l \equiv \sum_{l=1}^m U_l du_l,$$

at least, if this form in the u 's is not an exact differential. Let $b_{ij} = (\partial U_i / \partial u_j) - (\partial U_j / \partial u_i)$. We find from an elementary calculation that $b_{jk} = \sum a_{il} (\partial f_i / \partial u_j) (\partial f_l / \partial u_k)$. In other words, if we let A represent the matrix (a_{il}) , B the matrix (b_{jk}) and J the Jacobian matrix $(\partial f_i / \partial u_j)$, we have $B = J'AJ$, where J' is the transpose of J . The rank of J is m ; otherwise our manifold M would not be m -dimensional. It follows that the ranks of $J'A$ and AJ are each equal to m . Letting the rank of the skew symmetric matrix B be 2ν , we find from Frobenius' theorem† on the ranks of matrices that $\nu \geq m - k$. Thus, if $m > k = n/2$, the transformation S is sure to be Pfaffian. In any case, when $\nu > 0$, the parameters u_1, \dots, u_m can be "separated" just as the variables x_1, \dots, x_n were separated in (2.1). It is interesting to observe that B , and hence also ν , depends only on M and (1.1) and does not depend

* Whittaker [1, p. 275].

† MacDuffee [1, p. 11]. The theorem is stated for square matrices; but rectangular matrices can always be made square (without change of rank) by adding rows or columns of zeros.

upon the particular transformation T , so long as it is Pfaffian with respect to (1.1) and leaves M invariant.

The foregoing results are closely analogous to some results obtained recently by Wintner and van Kampen for differential equations.*

4. The characteristic multipliers. Let us now consider a nonsingular non-degenerate Pfaffian transformation in the neighborhood of an invariant point, which we take at the origin. In such a neighborhood T may be represented by power series

$$(4.1) \quad \bar{x}_i = \sum_{j=1}^{2k} c_{ij} x_j + \text{higher terms}.$$

Denoting by C the matrix of the c 's, we shall prove that *the latent roots of C occur in reciprocal pairs. That is, the $2k$ roots $\lambda_1, \lambda_2, \dots, \lambda_{2k}$ may be labelled so that $\lambda_{2i-1}\lambda_{2i} = 1$, ($i = 1, \dots, k$).*

Let α_{ij} be the constant term in the power series development of $a_{ij}(x)$. Then equating the constant terms in the identity (1.2), we obtain $\sum_{j,l=1}^{2k} \alpha_{jl} c_{ji} c_{lh} = \alpha_{ih}$. If we denote by A the matrix (α_{jl}) , these equations may be written in the form $C'AC = A$, where C' is the transpose of C . It follows that $A^{-1}C'A = C^{-1}$, where now the inverses A^{-1} and C^{-1} exist, since both A and C are nonsingular by hypothesis. Now corresponding to a latent root λ of C , we have a root λ^{-1} of C^{-1} occurring with the same multiplicity. On the other hand the above matrix equation shows that the latent roots of C^{-1} are the same as those of C' and hence also the same as those of C . The theorem readily follows, provided that we can prove that $\det C = +1$. Otherwise, there would be, for example, the possibility that C might admit two simple self-reciprocal roots $+1$ and -1 .

It is well known that a skew-symmetric determinant, typified by $\det A$, may be represented as follows:

$$2^k k! (\det A)^{1/2} = \sum_i \epsilon^{i_1 i_2 i_3 i_4 \dots i_{2k}} \alpha_{i_1 i_2} \alpha_{i_3 i_4} \dots \alpha_{i_{2k-1} i_{2k}},$$

where $\epsilon^{i_1 i_2 \dots i_{2k}}$ is $+1$ or -1 according as the permutation (i_1, \dots, i_{2k}) of the $2k$ integers $(1, 2, \dots, 2k)$ is even or odd. The summation \sum_i is extended over all permutations. But since $A = C'AC$, we have

$$\begin{aligned} 2^k k! (\det A)^{1/2} &= \sum_{j_1 \dots j_{2k}} \sum_i \epsilon^{i_1 i_2 \dots i_{2k}} \alpha_{j_1 j_2} \dots \alpha_{j_{2k-1} j_{2k}} c_{j_1 i_1} c_{j_2 i_2} \dots c_{j_{2k} i_{2k}} \\ &= (\det C) \sum_j \epsilon^{j_1 j_2 \dots j_{2k}} \alpha_{j_1 j_2} \dots \alpha_{j_{2k-1} j_{2k}} = (\det C) 2^k k! (\det A)^{1/2}. \end{aligned}$$

Here we use the fact that the only terms of the sum which do not cancel

* Wintner and van Kampen [1].

each other are the terms in which the j 's are mutually distinct. We also use the elementary definition of the value of a determinant. Since $\det A \neq 0$, the stated result that $\det C = +1$ follows at once.

The main result of this section is analogous to a corresponding classical result for differential equations, proved by Poincaré for the Hamiltonian case, and by G. D. Birkhoff and A. Wintner for the more general Pfaffian case.*

5. The formal differential system of a Pfaffian transformation. It is known† that, if S is an arbitrary real nonsingular analytic transformation, Pfaffian or not, which leaves the origin invariant, it is possible to find a positive integer N , such that the transformation S^N , which we shall hereafter call T , has the following properties:

I. T^t (transforming the point $x(0)$ into $x(t)$) may be represented in the neighborhood of the origin (for integral t) by means of power series of the form

$$(5.1) \quad x_i(t) = \sum_{j=1}^n y_{ij}(t)x_j(0) + \sum_{\mu=2}^{\infty} \left[\sum_{\alpha_1+\dots+\alpha_n=\mu} y_{i\alpha_1\dots\alpha_n}(t)x_1^{\alpha_1}(0)\dots x_n^{\alpha_n}(0) \right],$$

$i = 1, \dots, n,$

where the coefficients $y_{ij}, y_{i\alpha_1\dots\alpha_n}$, are entire functions of t , real when t is real, such that the series (5.1) converges when t is an integer. They do not in general converge for all t . The y 's also are such that any polynomial in a finite number of them, which vanishes for positive integral values of t , vanishes for all values of t .

II. There exist real formal power series $Y_i(x)$ in the variables x_1, \dots, x_n which are such that the right-hand members of (5.1) satisfy (in a formal sense) the formal differential equations

$$(5.2) \quad \frac{dx_i}{dt} = Y_i(x), \quad i = 1, \dots, n,$$

and reduce to $x_1(0), \dots, x_n(0)$, when $t=0$. It should be further noted that the series $Y_i(x)$ do not have constant terms.

We now proceed to study the formal differential system (5.2) in the case when S , and consequently T , is a Pfaffian transformation with respect to the form (1.1). We define the formal series $U_i(x)$ by means of the formulas

$$(5.3) \quad U_i(x) = \sum_{j=1}^n a_{ij}Y_j \equiv \sum_{j=1}^n \left(\frac{\partial X_i}{\partial x_j} - \frac{\partial X_j}{\partial x_i} \right) Y_j, \quad i = 1, \dots, n.$$

* Cf. Wintner [2].

† Lewis [1]. The integer N may, in general, be taken as unity.

We shall prove that *there exists a formal power series* $Q(x)$, *such that*

$$(5.4) \quad U_i(x) = \frac{\partial Q}{\partial x_i}.$$

Applying the identity (1.2) to the transformation T^t , for the case when t is an integer, we obtain

$$(5.5) \quad \sum_{i,l=1}^n a_{il}[x(t)] \frac{\partial x_i(t)}{\partial x_j(0)} \frac{\partial x_l(t)}{\partial x_h(0)} = a_{jh}[x(0)], \quad j, h = 1, \dots, n.$$

This identity stands for an infinite number of polynomial relations connecting the y 's which hold for integral values of t . Hence by property I, we easily see that (5.5) must hold formally for all values of t . Hence, writing x_i for $x_i(0)$, differentiating (5.5) formally with respect to t and then setting $t=0$, while remembering that $\partial x_i(0)/\partial x_h = \delta_{ih}$ and that $dx_i(t)/dt|_{t=0} = Y_i(x)$, we get the formal identities

$$(5.6) \quad \sum_{r=1}^n \frac{\partial a_{jh}}{\partial x_r} Y_r + \sum_{i=1}^n a_{ih} \frac{\partial Y_i}{\partial x_j} + \sum_{l=1}^n a_{jl} \frac{\partial Y_l}{\partial x_h} = 0.$$

From the definition of the a_{ij} first introduced in §1, we have $a_{ij} = -a_{ji}$ and

$$\frac{\partial a_{jh}}{\partial x_h} + \frac{\partial a_{hj}}{\partial x_h} + \frac{\partial a_{sh}}{\partial x_j} = 0.$$

With the help of these relations, (5.6) can readily be put into the form

$$\frac{\partial}{\partial x_h} \left[\sum_{s=1}^n a_{js} Y_s \right] = \frac{\partial}{\partial x_j} \left[\sum_{s=1}^n a_{hs} Y_s \right].$$

Hence referring back to (5.3), we see that we have established the identities $\partial U_i/\partial x_h = \partial U_h/\partial x_i$, ($j, h = 1, \dots, n$), from which the existence of a formal series Q satisfying (5.4) is obvious.

We now note that Q is *invariant under* T . For from (5.3) and (5.4) we find that

$$\sum_{i=1}^n Y_i \frac{\partial Q}{\partial x_i} = \sum_{i=1}^n Y_i U_i = \sum_{i,j=1}^n Y_i a_{ij} Y_j,$$

which vanishes identically in as much as $a_{ij} = -a_{ji}$. But the formal relation $\sum \partial Q/\partial x_i Y_i = 0$ is known to be necessary and sufficient that Q be invariant under T .

It is clear from (5.4), (5.3), and (5.2) that*

* The system (5.7) is equivalent to (5.2) in the nondegenerate case $2k=n$.

$$(5.7) \quad \sum_{j=1}^n a_{ij} \frac{dx_j}{dt} = \frac{\partial Q}{\partial x_i}, \quad i = 1, \dots, n.$$

We now ask ourselves what becomes of (5.7) if we make a change of variables. The answer is contained in the following theorem:

THEOREM. *Let the arbitrary nonsingular analytic transformation $x_i = x_i(y)$, ($i = 1, \dots, n$), carry the differential form $\sum_{i=1}^n X_i dx_i$ into $\sum_{i=1}^n Y_i dy_i$. For brevity, let $b_{ij} = (\partial Y_i / \partial y_j) - (\partial Y_j / \partial y_i)$. Then the formal differential equations (5.7) are carried over into $\sum_{j=1}^n b_{ij} dy_j / dt = \partial Q^* / \partial y_i$, ($i = 1, \dots, n$), where Q^* is the formal series in y_1, \dots, y_n obtained by substituting in Q the power series expressions for the x 's in terms of the y 's.*

In the sequel, when no confusion can occur, the asterisk will be omitted without further comment.

Proof. We note the elementary relations

$$b_{ij} = \sum_{h,l} a_{hl} \frac{\partial x_h}{\partial y_i} \frac{\partial x_l}{\partial y_j}.$$

Hence

$$\sum_{j=1}^n b_{ij} \frac{dy_j}{dt} = \sum_{h,l} a_{hl} \frac{\partial x_h}{\partial y_i} \frac{\partial x_l}{\partial y_j} \frac{dy_j}{dt} = \sum_{h,l} a_{hl} \frac{dx_l}{dt} \frac{\partial x_h}{\partial y_i},$$

and substituting from (5.7)[†] this becomes

$$\sum_{h=1}^n \frac{\partial Q}{\partial x_h} \frac{\partial x_h}{\partial y_i} = \frac{\partial Q}{\partial y_i}.$$

This completes the proof when once the obvious formal interpretation of the above symbols is supplied.[‡]

6. Canonical form of a Pfaffian transformation. It is well known[‡] that it is possible to choose coordinates, valid in the neighborhood of an assigned point, say the origin, in such a way that (1.1), or a differential form equivalent to it, may be written in the canonical form

$$(6.1) \quad \sum_{i=1}^k x_{2i} dx_{2i-1}.$$

Any transformation, Pfaffian with respect to (6.1), is said to be in *canonical form*. If it is nondegenerate ($2k = n$) and in canonical form, it is also called

[†] In case Q is convergent the invariance of the equations (5.7) is obvious from the fact that they express the conditions for the vanishing of the first variation of $\int (\sum X_i dx_i / dt + Q) dt$. But for Q divergent this simple proof needs a reinterpretation by no means obvious.

[‡] Weber [1, pp. 216-217].

a *contact transformation*. Any change of variables, which takes (6.1) into an equivalent form, will preserve the canonical form of a Pfaffian transformation by its very definition. Such a change of variables is itself a Pfaffian transformation in canonical form and is, therefore, also a contact transformation, if $2k = n$. On the other hand a change of variables which preserves the canonical form of a Pfaffian transformation need not be canonical nor even Pfaffian. It may, for example, change (6.1) into a constant multiple of itself.*

In case of a contact transformation, the formal differential equations (5.2) or (5.7) take the particularly simple Hamiltonian form

$$(6.2) \quad \frac{dx_{2i-1}}{dt} = -\frac{\partial Q}{\partial x_{2i}}, \quad \frac{dx_{2i}}{dt} = \frac{\partial Q}{\partial x_{2i-1}}, \quad i = 1, \dots, k.$$

Any change of variables which preserves the canonical form of the transformation will naturally preserve the Hamiltonian form of (5.2); but it is only when the transformation defining the change of variables is a contact transformation that the Hamiltonian of the transformed equations is the original Hamiltonian with the original variables replaced by their expressions in terms of the new variables. For other transformations further recourse must be had to the general theorem at the end of §5.

Let W be an arbitrary analytic function of x_{2i} and y_{2i-1} , ($i = 1, \dots, k$). Then the equations

$$(6.3) \quad y_{2i} = \frac{\partial W}{\partial y_{2i-1}}, \quad x_{2i-1} = \frac{\partial W}{\partial x_{2i}}, \quad i = 1, \dots, k,$$

are well known to define a contact transformation, provided that it is possible to solve (6.3) for the y 's in terms of the x 's and for the x 's in terms of the y 's. This is easily seen from the point of view of our present definition from the fact that

$$dW = \sum_{i=1}^k (y_{2i} dy_{2i-1} + x_{2i-1} dx_{2i})$$

is an exact differential; and hence also

$$dW - d\left(\sum_{i=1}^k x_{2i-1} x_{2i}\right) = \sum_{i=1}^k (y_{2i} dy_{2i-1} - x_{2i} dx_{2i-1})$$

is an exact differential. The transformations of this type, which leave the

* There are also transformations which may preserve the canonical character of particular contact transformations without having the general property. Thus any nonsingular transformation of coordinates preserves the canonical character of the identity transformation.

origin invariant, will play an important role in the sequel. Such transformations are known to form a group.

7. Normal form for the linear terms of a contact transformation in the case of simple elementary divisors. Suppose that a real contact transformation T , having the origin as an invariant point, is given by power series of the form (4.1). If the matrix (c_{ij}) has simple elementary divisors, it is known that it is possible to introduce a new set of variables (x'_1, \dots, x'_n) depending linearly on the original set (x_1, \dots, x_n) in such a way that T may be written in the normal form

$$(7.1) \quad \bar{x}'_j = \lambda_j x'_j + \text{higher terms}, \quad j = 1, \dots, 2k = n,$$

where $\lambda_1, \dots, \lambda_{2k}$ are the latent roots of (c_{ij}) . In accordance with the results of §4, the λ 's may be taken so that

$$(7.2) \quad \lambda_{2i-1} = e^{\rho_i}, \quad \lambda_{2i} = e^{-\rho_i}, \quad i = 1, \dots, k.$$

Simultaneously the differential form (6.1) will be transformed into the form

$$(7.3) \quad \sum_{i,j=1}^{2k} K_{ij} x'_i dx'_j,$$

and consequently T is no longer necessarily a contact transformation in the new variables.

The variables (x') are, moreover, in general no longer real; but they may be chosen in such a way that a real point in the original (x) -space always gives rise to a real value for x'_j , if λ_j is real, but to conjugate imaginary values for x'_r and x'_s , if λ_r and λ_s are conjugate imaginary. Variables of this type will be said to have the *property R*.

We now ask ourselves two questions: (I) Is it possible to make a linear change of variables in such a way that the linear terms of T are reduced to the form (7.1) while (7.3) is equivalent to a (nonzero) constant multiple of $\sum_{i=1}^k x'_{2i} dx'_{2i-1}$? In other words, can we bring it about that T , when expressed in normalizing variables, is still a contact transformation? (II) Is it possible to make the transformation satisfy the conditions of question (I) and in addition be such that the new variables have the property *R*?

The answer to question (I) is yes. The answer to question (II) is, in general, no, but is yes, if there are no conjugate imaginary λ 's of modulus unity, or if all the λ 's occur in conjugate imaginary pairs of modulus unity. However, even when the answer to (II) is in the negative, there always exist complex numbers $\alpha_1, \dots, \alpha_{2k}$ such that the $2k$ products $\alpha_1 x'_1, \dots, \alpha_{2k} x'_{2k}$ will have

the property R . To establish these statements the following discussion is sketched. The details are left to the reader.*

Suppose that we are already in possession of variables (x') reducing T to the normal form (7.1), transforming (6.1) to (7.3) and having the property R . We shall also at first assume that the λ 's are distinct. We proceed to replace (7.3) by the simplest equivalent differential form.

Since $K_{ii}x'_i dx'_i$ is an exact differential and since $K_{ij}x'_i dx'_j + K_{ji}x'_j dx'_i$ is equal to $(K_{ji} - K_{ij})x'_j dx'_i$ plus the exact differential $K_{ij}(x'_i dx'_j + x'_j dx'_i)$ and since, furthermore, in subtracting an exact differential from (7.3) we merely replace the latter by an equivalent form, we may assume that

$$(7.4) \quad K_{ji} = 0, \quad \text{whenever } j \leq i; \quad j, i = 1, 2, \dots, 2k.$$

Now the transformation given by (7.1) is Pfaffian with respect to (7.3). Hence, equating the constant terms in the identity corresponding to (1.2), we get

$$(K_{hi} - K_{ih})\lambda_i \lambda_h = K_{hi} - K_{ih}, \quad i, h = 1, \dots, 2k.$$

Hence $K_{hi} - K_{ih} = 0$ for all i and h for which $\lambda_i \lambda_h \neq 1$. If the λ 's are distinct, it follows from (7.2) and (7.4) that all the K_{ih} are zero except the $K_{2i, 2i-1}$, ($i = 1, \dots, k$). Thus we find that the differential form (7.3) is equivalent to

$$(7.5) \quad \sum_{i=1}^k M_i x'_{2i} dx'_{2i-1}, \quad M_i = K_{2i, 2i-1}.$$

None of these M_i can vanish since we are dealing with a nondegenerate Pfaffian transformation. We now consider the following two transformations for changing the variables in T :

$$(7.6) \quad x'_{2i-1} = x'_{2i-1}, \quad x'_{2i} = M_i x'_{2i},$$

$$(7.7) \quad x'_{2i-1} = M_i^{1/2} (-1)^{-1/4} x'_{2i-1}, \quad x'_{2i} = M_i^{1/2} (-1)^{-1/4} x'_{2i},$$

$$i = 1, \dots, k.$$

Neither of these transformations changes the form of the linear terms in (7.1), while the linear differential form (7.5) appears as $\sum_{i=1}^k x'_{2i} dx'_{2i-1}$ and $(-1)^{1/2} \sum_{i=1}^k x'_{2i} dx'_{2i-1}$.

In case none of the λ 's is in absolute value equal to 1, the notation can

* The reader will find it instructive to consider the various cases arising for $n=4$ with regard to the four characteristic multipliers $[\lambda_1, \lambda_2, \lambda_3, \lambda_4]$. If $\rho, \sigma, \theta, \phi$ are real, ($\rho, \sigma \neq 1$; $\theta, \phi \neq 0, \text{ mod } 2\pi$), the four important cases are as follows: I. $[\rho, \rho^{-1}, \sigma, \sigma^{-1}]$, II. $[\rho, \rho^{-1}, e^{\theta(-1)^{1/2}}, e^{-\theta(-1)^{1/2}}]$, III. $[e^{\theta(-1)^{1/2}}, e^{-\theta(-1)^{1/2}}, \rho, \rho^{-1}]$, IV. $[\rho e^{\theta(-1)^{1/2}}, \rho^{-1} e^{-\theta(-1)^{1/2}}, \rho e^{-\theta(-1)^{1/2}}, \rho^{-1} e^{\theta(-1)^{1/2}}]$. Case II is the one case in which canonical variables cannot be taken having the property R . It may be mentioned in this connection that the theorem in Birkhoff [3, p. 75, line 9] is incorrect as may be shown by simple examples.

be chosen in such a way that both members of any pair of imaginary roots will have indices with the same parity; so that, for example, λ_{2i} could be conjugate to λ_{2j} but not to λ_{2j-1} . If this convention is granted it can be shown that M_i is real, if λ_i is real, but is conjugate to M_j , if λ_i is conjugate to λ_j , and hence the variables (x'') have the property R .

If all the λ 's are equal in absolute value to 1, then each λ_{2i} is conjugate to its reciprocal λ_{2i-1} and to none other, since we are assuming the λ 's to be distinct. It can then be shown that M_i is pure imaginary. Let $M_i = N_i(-1)^{1/2}$. Without loss of generality we assume $N_i > 0$. Otherwise we interchange λ_{2i-1} and λ_{2i} , thus inducing the interchange x_{2i-1}'' and x_{2i}'' . In fact

$$x_{2i} dx_{2i-1} = -x_{2i-1} dx_{2i} + \text{the exact differential } d(x_{2i} x_{2i-1}).$$

With these changes made, if necessary, it is not difficult to show that the variables (x''') have the property R .

The case of equal latent roots (but simple elementary divisors) may be taken care of by a limiting process.

8. Normalization of Q in the formal differential equations. Let $T = S^N$, where S and N have the meanings explained at the beginning of §5. Suppose also that T is a contact transformation and that a preliminary normalization of the linear terms has been carried out as described in §7, so that T is Pfaffian with respect to $\sum_{i=1}^k x_{2i} dx_{2i-1}$ and is defined by convergent series of the type

$$(8.1) \quad \bar{x}_{2i-1} = e^{\rho_i} x_{2i-1} + \cdots, \quad \bar{x}_{2i} = e^{-\rho_i} x_{2i} + \cdots, \\ i = 1, 2, \cdots, k = n/2,$$

where only the linear terms have been written. It is assumed also that *there is a positive integer P (>3) such that there exist no relations of the form $\sum_{i=1}^k m_i \rho_i = 0$, where the integers m_1, \cdots, m_k are not all zero and where each m_i is numerically less than P* . Referring back to the formal Hamiltonian differential equations (6.2), let us write the formal series for $-Q$ in the form $-Q = \sum_{\alpha=2}^{\infty} H_{\alpha}$, where H_{α} is a homogeneous polynomial of the α th degree in x_1, \cdots, x_n . The formal series for T' have linear terms of the same form as those displayed in (8.1) except that ρ_i is replaced by $\rho_i t$. It follows at once from (6.2) that

$$H_2 = \sum_{i=1}^k \rho_i x_{2i-1} x_{2i}.$$

It is now our purpose to make changes in the variables with the help of contact transformations to see if H_3, H_4, \cdots can also be reduced to an especially simple form. Following Birkhoff's treatment of the analogous dynamical

problem,* we shall apply contact transformations of the form (6.3) with

$$W = \sum_{i=1}^m x_{2i} y_{2i-1} + W_\alpha, \quad \alpha = 3, 4, \dots,$$

where W_α is a homogeneous polynomial in the x_{2i} and y_{2i-1} of degree α . If we solve explicitly for the x 's in terms of the y 's we clearly obtain up to terms of the α th degree

$$x_{2i-1} = y_{2i-1} + \frac{\partial W'_\alpha}{\partial y_{2i}} + \dots, \quad x_{2i} = y_{2i} - \frac{\partial W'_\alpha}{\partial y_{2i-1}} + \dots, \quad i = 1, \dots, k.$$

Here W'_α denotes the function obtained by replacing x_{2i} by y_{2i} in W_α . To terms of the α th degree inclusive we find therefore that

$$-Q = \sum_{i=1}^k \rho_i y_{2i-1} y_{2i} + \sum_{i=1}^k \rho_i \left[y_{2i} \frac{\partial W'_\alpha}{\partial y_{2i}} - y_{2i-1} \frac{\partial W'_\alpha}{\partial y_{2i-1}} \right] + \sum_{s=3}^{\alpha} \bar{H}_s + \dots,$$

where \bar{H}_s is the polynomial obtained by replacing the x 's by the y 's in H_s .

In order to simplify the terminology let us now change the y 's into x 's. Thus, a contact transformation of this type leaves $H_2, H_3, \dots, H_{\alpha-1}$ unmodified while H_α takes the form

$$\sum_{i=1}^k \rho_i \left[x_{2i} \frac{\partial W''_\alpha}{\partial x_{2i}} - x_{2i-1} \frac{\partial W''_\alpha}{\partial x_{2i-1}} \right] + \text{the original } H_\alpha.$$

Here W''_α represents the function obtained by replacing y_{2i-1} by x_{2i-1} in W_α . Now any term in W''_α may be written $c x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}$, the β 's being positive integers whose sum is α . The corresponding term in the modified H_α has a coefficient

$$c \left[\sum_{i=1}^k \rho_i (\beta_{2i} - \beta_{2i-1}) \right] + h,$$

where h is the coefficient corresponding to c in the original H_α . Now, if $\alpha < P$, then $\sum_{i=1}^k (\beta_{2i} - \beta_{2i-1}) \rho_i$ cannot vanish according to the hypothesis italicized above, unless $\beta_{2i-1} = \beta_{2i}$ for each i , ($i = 1, 2, \dots, k$). Hence it is possible to choose the coefficients c in W''_α (or W_α) in such a way that all the terms in the modified H_α disappear except those which contain x_{2i-1} and x_{2i} both to the same power for each i , at least if $\alpha < P$.

Thus carrying out this process successively for $\alpha = 3, 4, \dots$, we see that after performing a finite number of contact transformations we may write

$$(8.2) \quad -Q = F(u) + Q_P,$$

* Birkhoff [3, pp. 82-85].

where $F(u)$ is a polynomial in the k products $x_{2i-1}x_{2i} = u_i$, ($i = 1, \dots, k$), of degree not greater than $P/2$, the linear terms being $\sum_{i=1}^k \rho_i u_i$, and where Q_P is a formal power series in x_1, \dots, x_n beginning with terms of degree not lower than P .

The sequence of contact transformations necessary to effect this reduction can, of course, be combined into a single analytic contact transformation. If there are no commensurability relations connecting ρ_1, \dots, ρ_k , P may, of course, be taken arbitrarily large. The process may then be continued indefinitely and we see that there exists a *formal* contact transformation which enables us to write Q as a formal power series in the u 's.

9. **Normalized forms for T .** The above normalization of Q yields immediately certain normalized forms for nondegenerate Pfaffian transformations. Namely, retaining the hypotheses and notation of the preceding section, T may be written in the form

$$(9.1) \quad \begin{aligned} \bar{x}_{2i-1} &= x_{2i-1} \exp [\partial F / \partial u_i] + \Xi_{2i-1}, \\ \bar{x}_{2i} &= x_{2i} \exp [-\partial F / \partial u_i] + \Xi_{2i}, \end{aligned} \quad i = 1, \dots, k = n/2,$$

where the Ξ 's are power series in the x 's beginning with terms of degree not lower than $P-1$. This is proved from the obvious fact that the formal differential equations determine the transformation T uniquely. Furthermore the power series Ξ_i must converge since the transformation T was analytic to begin with and only analytic changes of variable have been used. The remaining formal details of the proof are left to the reader.

The equations (9.1) are especially useful in case the characteristic exponents ρ_1, \dots, ρ_k are pure imaginary. In this case, if we start out with variables having the property R in the sense of §7, the variables subsequently introduced will also have this property. The proof of this fact is left to the reader. Transformations of this type have been considered by Birkhoff and Lewis.* In particular it is known that there exist infinitely many periodic point groups in the neighborhood of the origin, at least, if the Hessian determinant $|\partial^2 F / \partial u_i \partial u_j|$ evaluated at the origin is distinct from zero. It may be noted here that this result does not depend on complete incommensurability of the characteristic exponents, as assumed in the above mentioned work. It is merely necessary that P be not less than a certain fixed number which may be taken at least as small as $16k+10$.†

* Birkhoff [4]; Lewis [4, 5]. The transformation referred to appears in different notation at the top of page 119 of the first of these papers.

† In Birkhoff [4], it is merely assumed that $\mu \geq 8n+4$. The " n " of that paper is our k and " μ " = $P/2-1$.

REFERENCES

G. D. BIRKHOFF

1. *Surface transformations and their dynamical applications*, Acta Mathematica, vol. 43 (1922), pp. 1-119.
2. *Nouvelles recherches sur les systèmes dynamiques*, Memoriae Pontificiae Academiae Scientiae Novi Lyncae, (3), vol. 1, pp. 85-216.
3. *Dynamical Systems*, American Mathematical Society Colloquium Publications, vol. 9, New York, 1927.
4. With D. C. Lewis, *On the periodic motions near a given periodic motion of a dynamical system*, Annali di Matematica, (4), vol. 12 (1933), pp. 117-133.

E. GOURSAT

1. *Leçons sur le Problème de Pfaff*, Paris, 1922.

D. C. LEWIS

1. *On formal power series transformations*. To be published in the Duke Mathematical Journal.
2. *Invariant manifolds near an invariant point of unstable type*, American Journal of Mathematics, vol. 60 (1938), pp. 577-587.
3. *The formal theory of conservative transformations in 2n-dimensional space*, Bulletin of the American Mathematical Society, abstract 41-5-197.
4. *On certain periodic motions of dynamical systems with more than two degrees of freedom*, American Journal of Mathematics, vol. 56 (1934), pp. 25-41.
5. *Sulle oscillazioni periodiche d'un sistema dinamico*, Rendiconti della Reale Accademia Nazionale dei Lincei, (6*), vol. 19 (1934), pp. 234-237.

C. C. MACDUFFEE

1. *The Theory of Matrices*, Berlin, 1933.

E. VON WEBER

1. *Vorlesungen über das Pfaff'sche Problem und die Theorie der partiellen Differentialgleichungen erster Ordnung*, Leipzig, 1900.

E. T. WHITTAKER

1. *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 2d edition, Cambridge, 1917.

A. WINTNER

1. With E. R. van Kampen, *On the reduction of dynamical systems by means of parametrized invariant relations*, these Transactions, vol. 44 (1938), pp. 168-195.
2. *Three notes on characteristic exponents and equations of variation in celestial mechanics*, American Journal of Mathematics, vol. 53 (1931), pp. 605-625.

CORNELL UNIVERSITY,
ITHACA, NEW YORK

A STUDY OF CURVED SURFACES BY MEANS OF CERTAIN ASSOCIATED RULED SURFACES*

BY
P. O. BELL

Introduction. In this paper a point correspondence is introduced which is proving to be very helpful in the study of a general non-ruled analytic surface in ordinary space. If on a surface S tangents to the curves of an asymptotic family are constructed at the points of two curves of S which are not members of the family and which intersect in a point y of S , two ruled surfaces are thereby formed which have at y a common generator. The plane which is tangent to one of these ruled surfaces at a selected point of the common generator is tangent to the other at a distinct point whose location depends on the selection of the first point and on the choice of the two curves which determine the ruled surfaces. The use of this correspondence serves the following fourfold purpose: (1) to unify many of the apparently isolated topics which have been studied heretofore, (2) to interpret geometrically, by methods which are simpler than those formerly used, quantities which are intrinsically and projectively related to a surface, (3) to introduce and characterize new configurations which are covariantly related to a surface, and (4) to solve both recognized unsolved problems and new problems which present themselves in the theory.

1. Analytic basis. If the homogeneous projective coordinates $y^{(1)}, \dots, y^{(4)}$ of a general point y on a non-ruled surface S in ordinary space are analytic functions of two independent variables u, v , and if the parametric net on S is the asymptotic net, the functions $y^{(i)}$ are solutions of a system of differential equations, which by a suitable transformation can be reduced to Wilczynski's canonical form

$$(1.1) \quad y_{uu} + 2by_v + fy = 0, \quad y_{vv} + 2a'y_u + gy = 0.$$

The coefficients of these equations are functions of u, v , which are connected by three conditions of integrability.

In the notation employed by Green [1, p. 86], points ρ and σ on the u - and v -tangents to S at y , respectively, are given by

$$(1.2) \quad \rho = y_u - \beta y, \quad \sigma = y_v - \alpha y,$$

* Presented to the Society, September 6, 1938, under the title *Integral invariants of projective differential geometry*; §11 was presented April 9, 1937, under the title *Geometric characterizations in projective differential geometry of curved surfaces*; received by the editors May 19, 1939.

where α, β are functions of u, v . The line l joining the points ρ, σ generates a congruence Γ as y varies over S .

A line l' through a general point y of S , but not lying in the tangent plane to S at y , joins the point y to the point z given by

$$(1.3) \quad z = y_{uv} - \alpha y_u - \beta y_v,$$

wherein α, β are functions of u, v . As y varies over S the line l' generates a congruence Γ' . In accordance with the classification which Wilczynski introduced with his *directrices of the first and second kinds* [1, p. 95], we shall say that the line l and congruence Γ are of the first kind, and that the line l' and the congruence Γ' are of the second kind.* If the functions α, β are the same in equations (1.2), (1.3), the lines l, l' are called *reciprocal lines* because they are reciprocal polar lines with respect to the quadric of Lie at the point y . The congruences Γ, Γ' , generated by reciprocal lines, are called *reciprocal congruences*.

Throughout this paper, when no statement is made regarding the tetrahedron of reference for local coordinates of points or planes, the tetrahedron will be that whose vertices are y, y_u, y_v, y_{uv} . In this coordinate system, the equations for the line l are

$$(1.4) \quad x_1 + \beta x_2 + \alpha x_3 = 0, \quad x_4 = 0,$$

and the equations for l' are

$$(1.5) \quad x_2 + \alpha x_4 = 0, \quad x_3 + \beta x_4 = 0.$$

An arbitrary one-parameter family of curves on S is defined by the curvilinear equation

$$(1.6) \quad dv - \lambda du = 0,$$

where λ is an arbitrary function of u, v . We shall throughout this paper denote by F_λ the family defined by (1.6), and by C_λ the curve of the family which passes through the point y . The conjugate net N_λ of which F_λ is a family is defined by the curvilinear differential equation

$$(1.7) \quad dv^2 - \lambda^2 du^2 = 0.$$

We denote by C_λ and $C_{-\lambda}$ the two curves of N_λ which pass through the point y .

2. **The R_λ -associate of a line l .** As a point y of S moves along the curve C_λ , the u -tangent at y describes a ruled surface $R_\lambda^{(u)}$ and the v -tangent at y

* Green [1, p. 114] used this means of classifying his *canonical edges of the first and second kinds*. Fubini and Čech [1, pp. 96-102] have used the same means of classification but have reversed the names.

describes the ruled surface $R_\lambda^{(v)}$. The well known *asymptotic ruled surfaces* $R^{(u)}$ and $R^{(v)}$ are the special surfaces $R_\infty^{(u)}$ and $R_0^{(v)}$ in which the curves C_∞ and C_0 are the asymptotic v - and u -curves, respectively.

Since the ruled surfaces $R_\lambda^{(u)}$ and $R^{(u)}$ have at y the u -tangent to S as common generator, the plane which is tangent to $R_\lambda^{(u)}$ at a given point ρ of this generator is tangent to $R^{(u)}$ at another point ρ_λ of the generator. Likewise, since the ruled surfaces $R_\lambda^{(v)}$ and $R^{(v)}$ have at y the v -tangent to S as common generator, the plane which is tangent to $R_\lambda^{(v)}$ at a given point σ of this generator is tangent to $R^{(v)}$ at another point σ_λ of the generator. The points ρ_λ and σ_λ will be called the R_λ -transforms of the points ρ and σ , respectively. The line l_λ joining the points ρ_λ , σ_λ and the congruence Γ_λ generated by l_λ as y varies over S will be called the R_λ -associates of the line l and congruence Γ , respectively. The reciprocals of l_λ and Γ_λ will be called the R'_λ -associates of the reciprocals of l and Γ , respectively.

Let the point ρ_λ be given by $\rho_\lambda = y_u - \beta_\lambda y$, where β_λ is to be determined by the condition that the point $(\rho_\lambda)_v$ shall lie in the plane determined by the points y , ρ , and $\rho_u + \lambda \rho_v$. This condition is satisfied if, and only if, the line r_λ joining the points $\rho_u + \lambda \rho_v$, $(\rho_\lambda)_v$ cuts the corresponding u -tangent of S . By making use of equations (1.1) it may be shown that the line r_λ intersects the tangent plane to S at y in the point whose general coordinates are given by

$$\rho_u + \lambda \rho_v - \lambda (\rho_\lambda)_v = \lambda (\beta_\lambda - \beta - 2b/\lambda) y_v + f_1 y_u + f_2 y,$$

where f_1 and f_2 are nonzero functions of u, v . This point lies on the u -tangent if, and only if,

$$(2.1) \quad \beta_\lambda = \beta + 2b/\lambda.$$

In a similar manner the expression for the coordinates of σ_λ is found to be

$$(2.2) \quad \sigma_\lambda = y_v - \alpha_\lambda y,$$

where $\alpha_\lambda = \alpha + 2a'\lambda$.

3. The determination of the reference tetrahedra of Green. The general development of Green for the equation of a surface was referred to the tetrahedron whose vertices are points y , ρ , σ and τ , where τ is given by

$$(3.1) \quad \tau = y_{uv} - \alpha y_u - \beta y_v + \alpha \beta y$$

in which α, β are the same functions as those in (1.2) associated with the points ρ, σ . The point τ lies on the line l' which is the reciprocal of the line l joining ρ, σ . If the functions α, β are chosen suitably, the points ρ, σ and τ become covariant points, the coefficients in the development become absolute invariants, and the development is said to be a canonical development. Since

geometric determinations for the covariant points ρ, σ which are associated with the various canonical developments are well known, the completion of the problem of the determination of the covariant tetrahedra of Green is accomplished by the geometric characterization of the associated points τ .

Green [1, p. 98] has shown that the tangents at the points ρ, σ to the curved asymptotics of $R^{(u)}, R^{(v)}$, respectively, intersect in the point ω given by

$$(3.2) \quad \omega = \tau - 2a'by,$$

where τ corresponds to ρ, σ . The point ω_λ which is similarly defined with reference to the points $\rho_\lambda, \sigma_\lambda$ is given by $\omega_\lambda = \tau_\lambda - 2a'by$. By expressing the right member of this equation in terms of y, ρ, σ , and τ we have

$$(3.3) \quad \omega_\lambda = \tau + 2a'by - 2a'\lambda\rho - 2b\sigma/\lambda.$$

If the point y is kept fixed, while the function λ is varied, the line joining y and ω_λ describes a quadric cone whose equation when referred to y, ρ, σ, τ is found to be

$$(3.4) \quad x_2x_3 - 4a'bx_4^2 = 0.$$

Moreover, since for every value of λ the expression ω_λ is a linear combination of $\tau + 2a'by, \rho$ and σ , the points ω_λ all lie in the plane π_ω determined by these three points. The locus of the points ω_λ as λ is varied is therefore a conic. The point $\tau + 2a'by$ is the intersection of the plane π_ω with the line l' . Finally, the point τ is the harmonic conjugate of y with respect to the points $\tau + 2a'by$ and $\tau - 2a'by$.

We find also that the cone (3.4) intersects the tangent plane in the asymptotic tangents to S at y . The planes which are tangent to the cone along the u - and v -tangents to S at y intersect in the line l' which is the reciprocal of the line joining ρ, σ .

4. The family of R_λ -derived curves and the curves of Darboux and of Segre. Let Q_λ denote the point of intersection of a line l with its R_λ -associate l_λ . The point Q_λ is given by

$$(4.1) \quad Q_\lambda = a'\lambda^2\rho - b\sigma.$$

The direction of the tangent to S at y joining the points y, Q_λ is given by

$$(4.2) \quad dv/du = -b/a'\lambda^2$$

where the right member is evaluated at the point y . The tangent line in this direction will be called the R_λ -correspondent of the tangent to the curve C_λ at y . The one-parameter family of curves defined by the curvilinear differential equation

$$(4.3) \quad a'\lambda^2dv + bdu = 0$$

will be called the family of R_λ -derived curves. This family is completely characterized by the property that at a general point y of S the tangent to the R_λ -derived curve through y is the R_λ -correspondent of the tangent to the curve C_λ which passes through the point. We observe that the family of R_λ -derived curves is independent of the choice of the congruence Γ used in the definition and is the same as the family of $R_{-\lambda}$ -derived curves. Hence, we may associate the R_λ -derived curves with a conjugate net.

Since the curvilinear differential equation for the curves of Darboux is

$$(4.4) \quad a'dv^3 + bdu^3 = 0,$$

and that for the curves of Segre is

$$(4.5) \quad a'dv^3 - bdu^3 = 0,$$

the following theorems are immediate consequences.

THEOREM 4.1. *A curve C_λ is a curve of Darboux if, and only if, at each of its points the R_λ -correspondent of the tangent to C_λ coincides with this tangent.*

THEOREM 4.2. *A curve C_λ is a curve of Segre if, and only if, at each of its points the tangent to C_λ and its R_λ -correspondent are conjugate tangents of S .*

5. Pencils of conjugate nets; the Segre-Darboux pencil. The class of ∞^1 conjugate nets on S , every one of which has the property that at every point of the surface its two tangents form with the tangents of a fundamental conjugate net the same cross ratio, is called a pencil of conjugate nets (Wilczynski [2, p. 216]). The differential equation of a general net $N_{\lambda\lambda_1}$ of the pencil p_{λ_1} of conjugate nets determined by the fundamental net N_{λ_1} , defined by $dv^2 - \lambda_1^2 du^2 = 0$, is of the form

$$(5.1) \quad dv^2 - h^2 \lambda_1^2 du^2 = 0, \quad (h = \text{const.}).$$

DEFINITION 5.1. *The conjugate net N_{λ_1} , where $\lambda_1 = -b/a'\lambda^2$ will be called the R_λ -derived conjugate net associated with the family F_λ defined by (1.6). The associated pencil p_{λ_1} will be called the R_λ -derived pencil of conjugate nets.*

The curves of Darboux and the curves of Segre belong to a pencil of conjugate nets called the Segre-Darboux pencil. The curvilinear differential equation for this pencil is (5.1), where $\lambda_1 = (b/a')^{1/3}$. As an immediate consequence of the form of this equation we have

THEOREM 5.1. *If a conjugate net N_λ is contained in the associated R_λ -derived pencil of conjugate nets, the pencil is the Segre-Darboux pencil.*

The following theorem presents an additional characteristic property of this pencil.

THEOREM 5.2. *A conjugate net N_λ belongs to the Segre-Darboux pencil of conjugate nets if, and only if, at a general point y of S , its axis lies in the osculating plane of the R_λ -derived curve at this point.*

Let $C_{\bar{\lambda}}$ denote the R_λ -derived curve which passes through the point y . The direction of $C_{\bar{\lambda}}$ at y is given by $dv - \bar{\lambda}du = 0$, where $\bar{\lambda} = -b/a'\lambda^2$. The equation of the osculating plane of $C_{\bar{\lambda}}$ at y is

$$(5.2) \quad 2\bar{\lambda}^2x_2 - 2\bar{\lambda}x_3 + (\bar{\lambda}' - 2b + 2a'\bar{\lambda}^3)x_4 = 0.$$

The axis at y of the conjugate net N_λ is the line joining the point y and the point z given by (1.3), wherein

$$\alpha = (\log \lambda)_v/2 - b/\lambda^2, \quad \beta = -(\log \lambda)_u/2 - a'\lambda^2.$$

The axis of N_λ at y lies in the plane (5.2) if the local coordinates $(0, -\alpha, -\beta, 1)$ of the point z satisfy equation (5.2). It is convenient to express this condition in terms of $\bar{\lambda}$. Making use of the following relations,

$$\begin{aligned} (\log \lambda)_v/2 - b/\lambda^2 &= (\log kb/a'\bar{\lambda})_v/4 + a'\bar{\lambda}, \\ -(\log \lambda)_u/2 - a'\lambda^2 &= -(\log kb/a'\bar{\lambda})_u/4 + b/\bar{\lambda}, \end{aligned} \quad (k = \text{const.}),$$

we obtain this condition in the form

$$2\bar{\lambda}' = \bar{\lambda}([\log kb/a'\bar{\lambda}]_u + [\log kb/a'\bar{\lambda}]_v\bar{\lambda}),$$

which may be reduced to the simpler form

$$(5.3) \quad 2[\log \bar{\lambda}]' = [\log kb/a'\bar{\lambda}]',$$

in which accents indicate differentiation with respect to the independent variable u , and $dv/du = \bar{\lambda}$. On integrating (5.3) we obtain

$$\log \bar{\lambda}^2 = \log (kb/a'\bar{\lambda}), \quad (k = \text{arb. const.}).$$

Hence, we have

$$(5.4) \quad \bar{\lambda} = \epsilon(kb/a')^{1/3}, \quad (\epsilon^3 = 1).$$

Solving for λ , making use of the relation $\bar{\lambda}^2 = -b/a'\lambda^2$, we obtain

$$(5.5) \quad \lambda = \pm i\epsilon(1/k)^{1/6}(b/a')^{1/3}, \quad (i = (-1)^{1/2}).$$

Since k is an arbitrary constant, the net N_λ , where λ is given by (5.5), belongs to the Segre-Darboux pencil. This establishes the sufficiency of the condition. The condition can be shown to be necessary by interchanging the hypothesis and conclusion of the sufficiency proof, and reversing the argument.

The conjugate nets of simplest description, which have the property described in this theorem, are the Segre-Darboux nets. Without the foregoing argument it is clear that the Segre-Darboux nets have the property, since from the theorems of §5 we have that *the R_{λ_i} -derived family of a Segre-Darboux net N_{λ_i} , ($i=1, 2, 3$), is the family of Darboux curves of the net.*

6. Projective characterizations for the Γ -curves of a congruence Γ . Green defined the Γ -curves of the congruence Γ to be the curves of S which correspond to the developables of the congruence Γ . New projective characterizations for these curves will be presented in this section.

As y varies over the surface S , the points ρ and σ of a line of the first kind generate transversal surfaces S_ρ and S_σ of a congruence Γ . Since corresponding points y, ρ and σ have the same curvilinear coordinates (u, v) , corresponding directions at these points are defined by the same ratio $dv/du = \lambda$ and the correspondences between pencils of tangents at y , at ρ , and at σ , are projectivities. Let us denote by π_y, π_ρ and π_σ the planes which are tangent to S, S_ρ , and S_σ at y, ρ , and σ , respectively. For an unspecialized surface S the tangent planes π_ρ and π_σ intersect in a line h , and the two projective pencils of tangents at ρ and σ determine a projectivity on h which has two distinct double points P_1 and P_2 . The two directions λ_1, λ_2 which correspond to these double points are therefore (for an unspecialized surface) the only ones for which the points $\rho, \rho_u + \lambda \rho_v, \sigma_u + \lambda \sigma_v$ are coplanar. The condition that these points be coplanar is necessary and sufficient that for this direction the line l joining ρ, σ describes a developable surface of Γ . Hence, λ_1 and λ_2 are the directions of the Γ -curves of S at y . Moreover, if we consider the definition of the R_λ -associate, it is clear that the reciprocals $l_{\lambda_1}, l_{\lambda_2}$ of the lines joining y and P_1 , and y and P_2 , respectively, are the R_{λ_1} - and R_{λ_2} -associates of l . Hence, the lines joining y and P_1 , and y and P_2 , are the R_{λ_1}' - and R_{λ_2}' -associates, respectively, of the line l' which is the reciprocal of l . We have, therefore, the following theorem:

THEOREM 6.1. *If, and only if, at each point y of a curve C_λ of S , the line of the R_λ' -associate of the reciprocal congruence Γ' passes through the line h of intersection of the tangent planes to S_ρ at ρ and S_σ at σ , the curve is a Γ -curve of the congruence Γ .*

By considering polar reciprocals with respect to a quadric of Darboux, we obtain

THEOREM 6.2. *A curve C_λ of S is a Γ -curve of a congruence Γ if, and only if, at each point y of the curve the pole of the plane π_y , determined by the point y and the line of intersection of the tangent planes to S_ρ at ρ and S_σ at σ , with respect to a quadric of Darboux lies on the R_λ -associate of the line of the congruence Γ which corresponds to the point y .*

The curvilinear differential equation for the net of Γ -curves of S may be put in the form

$$(6.1) \quad 2b(\alpha - \alpha_{(a)})du^2 + (\beta_v - \alpha_u)dvd u - 2a'(\beta - \beta_{(a)})dv^2 = 0,$$

wherein $\alpha_{(a)} = -(f + \beta_u + \beta^2)/2b$, $\beta_{(a)} = -(g + \alpha_v + \alpha^2)/2a'$. The equations for the planes π_p and π_σ may be shown to be

$$(6.2) \quad x_1 + \beta x_2 + \alpha_{(a)}x_3 + (\beta\alpha_{(a)} + \beta_v)x_4 = 0,$$

$$(6.3) \quad x_1 + \beta_{(a)}x_2 + \alpha x_3 + (\alpha\beta_{(a)} + \alpha_u)x_4 = 0,$$

respectively. By making use of these equations in connection with equation (6.1) for the Γ -curves, the following theorems may be easily proved.

THEOREM 6.3. *The planes π_p and π_σ intersect the line l' which is the reciprocal of the line joining ρ , σ in one and the same point if, and only if, $\beta_v = \alpha_u$.*

THEOREM 6.4. *If the planes π_p and π_σ intersect in a line h which contains a point of l' , the Γ -curves form a conjugate net when h passes through neither asymptotic tangent of S at y , but they coincide with the family of asymptotic v -curves (or u -curves) of S when h passes through the asymptotic u -tangent (or v -tangent) to S at y .*

THEOREM 6.5. *If, and only if, the plane π_p coincides with the plane π_σ , the Γ -curves are indeterminate.*

THEOREM 6.6. *If $\beta_v \neq \alpha_u$, the planes π_p and π_σ intersect in a line h which is not coplanar with l' . If the line h contains the point ρ (or σ), one family of the net of Γ -curves is the family of asymptotic v -curves (or u -curves). If h coincides with the line l , the net of Γ -curves coincides with the asymptotic net on S .*

DEFINITION 6.1. *A congruence which satisfies the hypothesis of Theorem 6.3 will be called central to S .*

DEFINITION 6.2. *Let $\sigma_{(a)}$ denote the intersection of the plane π_p with the tangent to the asymptotic v -curve of S at y , and let $\rho_{(a)}$ denote the intersection of the plane π_σ with the tangent to the asymptotic u -curve of S at y . The line $l_{(a)}$ joining $\rho_{(a)}$, $\sigma_{(a)}$ and the congruence $\Gamma_{(a)}$ generated by $l_{(a)}$ as y varies over S will be called the asymptotic associates of the line l and congruence Γ respectively.*

The points $\rho_{(a)}$, $\sigma_{(a)}$ * are given by

$$(6.4) \quad \rho_{(a)} = y_u - \beta_{(a)}y, \quad \sigma_{(a)} = y_v - \alpha_{(a)}y,$$

where $\beta_{(a)}$, $\alpha_{(a)}$ are the functions which appear in (6.1).

* The points $\rho_{(a)}$, $\sigma_{(a)}$ may also be characterized as follows. The point $\rho_{(a)}$ is the intersection of the v -tangent of S_σ at σ with the u -tangent of S at y , and the point $\sigma_{(a)}$ is the intersection of the u -tangent of S_ρ at ρ with the v -tangent of S at y .

The Γ' -curves are the curves of S which correspond to the developables of the congruence Γ' ; they are in fact the intersections of the surface S with the developables of the congruence Γ' .

DEFINITION 6.3. *If one of the two families of curves which form a conjugate net consists of Γ' -curves, we shall call the other a family of reflected Γ' -curves.*

Green [1, p. 93] defined reflected Γ -curves in a similar manner.

The curvilinear differential equations for the net of Γ' -curves of S may be put in the form

$$(6.5) \quad (2b_v - 2b[\alpha + \alpha_{(a)}])du^2 + (\beta_v - \alpha_u)dudv + (2a'[\beta + \beta_{(a)}] - 2a'_u)dv^2 = 0.$$

The equation for the net of reflected Γ' -curves is

$$(6.6) \quad [2b(\alpha + \alpha_{(a)}) - 2b_v]du^2 + (\beta_v - \alpha_u)dudv + [2a'_u - 2a'(\beta + \beta_{(a)})]dv^2 = 0.$$

The Γ -curves coincide with the reflected Γ' -curves if, and only if

$$\alpha_{(a)} = b_v/2b, \quad \beta_{(a)} = a'_u/2a'.$$

For this selection the line $l_{(a)}$ is the directrix of the first kind of Wilczynski. Hence we have

THEOREM 6.7. *The Γ -curves of a congruence Γ coincide with the reflected Γ' -curves of the reciprocal congruence Γ' if, and only if, the asymptotic associate of the congruence Γ is the congruence generated by the directrix of the first kind of Wilczynski.*

The following theorem may be proved similarly. Let $P_{(a)}$ denote the point of intersection of the line l with its asymptotic associate $l_{(a)}$. Let $t_{(a)}$ denote the tangent to S at y which passes through $P_{(a)}$, and let $t'_{(a)}$ denote the tangent to S at y which is conjugate to $t_{(a)}$ at y .

THEOREM 6.8. *The Γ' -curves of a congruence Γ' are indeterminate if, and only if, the following conditions are satisfied: (1) the congruence Γ is central to S , (2) the pencil of lines whose center is the point $P_{(a)}$ contains the directrix of the first kind of Wilczynski, and (3) the tangent line $t_{(a)}$ and the directrix of the first kind, separate harmonically the lines l and $l_{(a)}$.*

The second and third conditions in the theorem may be replaced by the following ones: (2') the pencil of lines determined by the lines l' and $l'_{(a)}$ contains the directrix of the second kind of Wilczynski, as well as the tangent $t'_{(a)}$ which is conjugate to $t_{(a)}$, (3') the tangent line $t'_{(a)}$ and the directrix of the second kind separate harmonically the lines l' and $l'_{(a)}$.

7. **Theorems on conjugate nets.** According to Theorems 6.4 and 6.5 the congruences central to S consist of (1) congruences harmonic to S , (2) congruences

whose Γ -curves coincide with an asymptotic family of S , and (3) congruences whose Γ -curves are indeterminate.

DEFINITION 7.1. A family F_λ defined by (1.6) belongs to class \mathfrak{C} if the R_λ -associate of a congruence central to S is likewise central to S .

The analytic condition that F_λ belong to class \mathfrak{C} is that λ satisfy the partial differential equation

$$(7.1) \quad (b/\lambda)_v - (a'\lambda)_u = 0.$$

If a direction $dv/du = \lambda$ satisfies equation (7.1), its conjugate direction $dv/du = -\lambda$ likewise satisfies it. Hence, we have that if a family belongs to class \mathfrak{C} , the conjugate net N_λ , of which F_λ is a family, belongs to class \mathfrak{C} .

THEOREM 7.1. If a family F_λ of curves of S belongs to class \mathfrak{C} , the R_λ -derived conjugate net consists of a one-parameter family of projective geodesics and the family of R_λ -derived curves.

The function λ satisfies equation (7.1) which may be put in the form

$$(7.2) \quad \bar{\lambda}(\log \lambda)_v + (\log \lambda)_u = -(\log a')_u + \bar{\lambda}(\log b)_v,$$

wherein $\bar{\lambda} = b/a'\lambda^2$. We obtain from the equation $\lambda^2 = b/a'\bar{\lambda}$ by logarithmic differentiation, the relations

$$(7.3) \quad \begin{aligned} (\log \lambda)_v &= [(\log b/a')_v - (\log \bar{\lambda})_v]/2, \\ (\log \lambda)_u &= [(\log b/a')_u - (\log \bar{\lambda})_u]/2. \end{aligned}$$

Using these forms, we may express equation (7.2) entirely in terms of $\bar{\lambda}$. On simplifying the resulting equation we obtain

$$(7.4) \quad \bar{\lambda}_u + \bar{\lambda}\bar{\lambda}_v = (\log a'b)_u\bar{\lambda} - (\log a'b)_v\bar{\lambda}^2.$$

Putting $\bar{\lambda} = dv/du$ and $\bar{\lambda}_u + \bar{\lambda}\bar{\lambda}_v = d^2v/du^2$, we have the usual form of the differential equation for the projective geodesics

$$(7.5) \quad d^2v/du^2 = (\log a'b)_u dv/du - (\log a'b)_v (dv/du)^2.$$

Hence the curves defined by $dv - \bar{\lambda}du = 0$, where $\bar{\lambda} = b/a'\lambda^2$, are projective geodesics under the hypothesis of the theorem. Since, moreover, the family of R_λ -derived curves is defined by $dv - \lambda_1 du = 0$, where $\lambda_1 = -b/a'\lambda^2$, the families $F_{\bar{\lambda}}$ and F_{λ_1} form the R_λ -derived conjugate net.

THEOREM 7.2. If a one-parameter family of projective geodesics and the family of R_λ -derived curves associated with a family F_λ form a conjugate net, the family F_λ belongs to class \mathfrak{C} .

A one-parameter family of projective geodesics is defined by $dv - \bar{\lambda} du = 0$, where $\bar{\lambda}$ is a solution of the equation (7.5). According to hypothesis, $\bar{\lambda} = b/a'\lambda^2$. Hence, we have

$$(7.6) \quad \bar{\lambda}_u + \bar{\lambda}\bar{\lambda}_v = (b/a'\lambda^2)_u + (b/a'\lambda^2)(b/a'\lambda^2)_v.$$

Equating the right members of (7.4) and (7.6) and simplifying, we obtain

$$(7.7) \quad (\log a'\lambda)_u = (b/a'\lambda^2)(\log b/\lambda)_v,$$

which is equivalent to equation (7.1).

The functions α, β for the axis congruence and the associate axis congruence of the net N_λ are given by

$$(7.8) \quad \alpha = (\log \lambda)_v/2 - b/\lambda^2, \quad \beta = -(\log \lambda)_u/2 - a'\lambda^2,$$

$$(7.9) \quad \alpha = (\log \lambda)_v/2 + b/\lambda^2, \quad \beta = -(\log \lambda)_u/2 + a'\lambda^2,$$

respectively. Hence, we have

THEOREM 7.3. *If F^- denotes the family of R_λ -derived curves associated with a conjugate net N_λ , the axis congruence of the net is the R_λ^- -associate of the associate axis congruence of the net.*

Since, moreover, the ray congruence of a conjugate net is the reciprocal of the associate axis congruence of the net, and the associate ray congruence is the reciprocal of the axis congruence, we have

THEOREM 7.4. *If F_λ^- denotes the family of R_λ -derived curves associated with a conjugate net N_λ , the associate ray congruence of the net is the R_λ^- -associate of the ray congruence of the net.*

8. The $R_{\lambda,j,k}$ -associates of a line l and congruence Γ ; the transformations of Čech. The concept of the R_λ -associate of a line l may be generalized as follows. Let $\rho_{\lambda,j}$ denote the point on the asymptotic u -tangent to S at y which is determined by the cross ratio equation

$$(8.1) \quad (y, \rho, \rho_\lambda, \rho_{\lambda,j}) = j, \quad (j = \text{const.}).$$

Let $\sigma_{\lambda,k}$ denote the point determined on the v -tangent to S at y by the cross ratio equation

$$(8.2) \quad (y, \sigma, \sigma_\lambda, \sigma_{\lambda,k}) = k, \quad (k = \text{const.}).$$

The points $\rho_{\lambda,j}$ and $\sigma_{\lambda,k}$ will be called the $R_{\lambda,j}$ - and $R_{\lambda,k}$ -transforms of ρ and σ , respectively. The line $l_{\lambda,j,k}$ joining the points $\rho_{\lambda,j}$ and $\sigma_{\lambda,k}$ and the congruence $\Gamma_{\lambda,j,k}$ generated by $l_{\lambda,j,k}$ as y moves over S will be called the $R_{\lambda,j,k}$ -associates of l and Γ respectively. The reciprocals $l'_{\lambda,j,k}$ and $\Gamma'_{\lambda,j,k}$ of $l_{\lambda,j,k}$ and $\Gamma_{\lambda,j,k}$ will be called the $R'_{\lambda,j,k}$ -associates of l' and Γ' , respectively.

The points $\rho_{\lambda,j}$ and $\sigma_{\lambda,k}$ are given by

$$(8.3) \quad \rho_{\lambda,j} = \rho - 2jby/\lambda, \quad \sigma_{\lambda,k} = \sigma - 2ka'\lambda y.$$

The general transformation $\Sigma_{j,k}$ of Čech* [1, p. 192] is defined analytically by the equations

$$(8.4) \quad \begin{aligned} \xi_1 &= 0, & r\xi_2 &= x_2x_3^2, & r\xi_3 &= x_2^2x_3, \\ r\xi_4 &= -x_1x_2x_3 - 2bjx_2^3 - 2a'kx_3^3, \end{aligned} \quad (j = \text{const.}, k = \text{const.}),$$

where r is a proportionality factor. It is a transformation between points with local coordinates x in the tangent plane of a surface at a point and planes with local coordinates ξ through the point.

We present a new geometric characterization of the general transformation of Čech. Let t_λ and $t_{-\lambda}$ denote the conjugate tangents to S at y whose directions are $dv/du = \lambda$, and $dv/du = -\lambda$, respectively. Let $P_{-\lambda}$ denote the point of intersection of $t_{-\lambda}$ with an arbitrary line l of the first kind.

THEOREM 8.1. *The plane $\pi_{\lambda,j,k}$ which corresponds to a point $P_{-\lambda}$ in the transformation $\Sigma_{j,k}$ of Čech is the plane which is determined by the tangent t_λ and the reciprocal of the $R_{\lambda,j,k}$ -associate of l .*

The equations for $t_{-\lambda}$ are $x_3 + \lambda x_2 = 0$, $x_4 = 0$. Equations (1.4) are for an arbitrary line l . Hence the local coordinates of $P_{-\lambda}$ may be found to be

$$(8.5) \quad x_1 = -\beta + \alpha\lambda, \quad x_2 = 1, \quad x_3 = -\lambda, \quad x_4 = 0.$$

There is a point Q_λ of t_λ whose local coordinates are given by $(0, 1, \lambda, 0)$, and there is a point $z_{\lambda,j,k}$ of $l'_{\lambda,j,k}$ whose local coordinates are $(0, -\alpha - 2ka'\lambda, -\beta - 2jb/\lambda, 1)$. Since the plane determined by t_λ and $l'_{\lambda,j,k}$ contains the points y , Q_λ , and $z_{\lambda,j,k}$, its equation may be found to be

$$(8.6) \quad \lambda x_2 - x_3 + (\alpha\lambda + 2ka'\lambda^2 - \beta - 2jb/\lambda)x_4 = 0.$$

By substituting the values for x_1, x_2, x_3, x_4 given by (8.5) in (8.4), we find that the coordinates of the plane $\pi_{\lambda,j,k}$ are the same as those of the plane (8.6), except for a proportionality factor. Hence the theorem is proved.

The correspondence of C. Segre is the transformation $\Sigma_{1,1}$. It was defined to be the correspondence between the osculating planes at a point y of S of all of the curves of S passing through y and the corresponding ray-points of these curves at y . The geometrical characterization which we have given for $\Sigma_{j,k}$ reduces to a very simple form for $\Sigma_{1,1}$, namely, *the plane π_λ which is in the correspondence of Segre with the point $P_{-\lambda}$ is determined by the tangent t_λ and the line l'_λ which is the reciprocal of the R_λ -associate of l .*

* Lane [1, p. 209] has characterized geometrically the transformations $\Sigma_{j,k}$, where $j=k$.

To locate the ray-point of the curve C_λ at y , select in the osculating plane π_λ , of the curve C_λ at y , an arbitrary line l' of the second kind. The $R_{-\lambda}$ -associate of the reciprocal of l' intersects the tangent line $t_{-\lambda}$ in the point $P_{-\lambda}$ which is the ray-point of C_λ at y .

The equations in plane coordinates ξ of the pencil of planes whose axis is an arbitrary line l' of the second kind are $\xi_1=0$, $\xi_4-\alpha\xi_2-\beta\xi_3=0$, wherein α, β are arbitrary functions of u, v . It is well known that the points which correspond to these planes in the transformation $\Sigma_{j,k}$ determine a plane cubic in the tangent plane to S at y whose equations are

$$(8.7) \quad x_1x_2x_3 + \alpha x_2x_3^2 + \beta x_2^2x_3 + 2jb x_2^3 + 2ka'x_3^3 = 0, \quad x_4 = 0.$$

Let $K_{\lambda,j,k}$ denote the point of intersection of the tangent t_λ with the $R_{\lambda,j,k}$ -associate of l .

THEOREM 8.2. *The locus of the point $K_{\lambda,j,k}$ as the direction λ is varied at y is the cubic (8.7).*

The equations for t_λ and for $l_{\lambda,j,k}$ are $x_3=\lambda x_2$ and $x_1+(\beta+2jb/\lambda)x_2+(\alpha+2ka'\lambda)x_3=0$, respectively. The point $K_{\lambda,j,k}$ of intersection of these two lines has, therefore, coordinates that are proportional to

$$(8.8) \quad x_1 = -(\beta + 2jb/\lambda + \alpha\lambda + 2ka'\lambda^2), \quad x_2 = 1, \quad x_3 = \lambda.$$

Homogeneous elimination of λ among equations (8.8) gives the equation (8.7) for the cubic.

Among the important cubics (8.7) which have been studied is the one introduced by B. Segre in which all of the ∞^4 non-composite cubic surfaces having fourth order contact with the surface at the point y cut the tangent plane of the surface at y . This cubic is characterized geometrically by Theorem 8.2, wherein $j=k=1/3$ and l is the canonical edge of the first kind. Its equations in the notation of this paper are equations (8.7), wherein $j=k=1/3$ and $\alpha=a_u'/4a'$, $\beta=b_u/4b$.

9. Differential invariants. The differential form

$$(9.1) \quad d\phi_{j,k} = -2(jbdu^3 + ka'dv^3)/dudv, \quad (j, k = \text{const.}),$$

is an absolute invariant under the most general transformation of independent and dependent variables maintaining the asymptotic net as parametric.

To provide a geometric interpretation for $d\phi_{j,k}$ let Q denote a point on the tangent to a curve C_λ at y and let K and $K_{\lambda,j,k}$ denote the points in which this tangent intersects a line l and its $R_{\lambda,j,k}$ -associate line, respectively. We define the (j, k) non-euclidean distance from y to Q to be

$$(9.2) \quad D_{yQ}^{(j,k)} = (y, K, Q, K_{\lambda,j,k}).$$

Let Y denote a point near to y on the curve C_λ , and let the curvilinear coordinates of y and Y be (u, v) and $(u + \delta u, v + \delta v)$ respectively, where $(\delta u^2 + \delta v^2)^{1/2} \leq \epsilon$. Since $Y = y(u + \delta u, v + \delta v)$ and the limit of $\delta v / \delta u$ as δu tends to zero as $\lambda(u, v)$, the general coordinates of Y may be given by the expansion

$$Y = y + (y_u + \lambda y_v) \delta u + \text{terms of order } (\delta u)^2$$

wherein $y = y(u, v)$. Hence the coordinates of Y differ only by terms of order $(\delta u)^2$ from the point Y_1 on the tangent to C_λ at y given by $Y_1 = y + (y_u + \lambda y_v) \delta u$. Therefore the principal parts of the infinitesimal cross ratios

$$(y, K, Y, K_{\lambda, j, k}), \quad (y, K, Y_1, K_{\lambda, j, k})$$

are identical. It may be easily shown that this principal part is the absolute differential invariant $d\phi_{j, k}$ which we wished to characterize.

It may be observed that $d\phi_{1, 1}$ is the *projective linear element* and $d\phi_{0, 1}$ and $d\phi_{1, 0}$ are the *elementary forms* of Bompiani.

The integral of the form $d\phi_{j, k}$ extended over a finite arc C_λ is intrinsically and projectively related to this arc. To interpret this integral geometrically let A and B denote the end points of the arc, and let (u_0, v_0) and (u, v) denote the curvilinear coordinates of A and B , respectively. Let ϵ be a positive number, and divide the arc C_λ by means of the intermediate points Y_i , ($i = 1, 2, \dots, n-1$), into n smaller arcs. Let the curvilinear coordinates of Y_p be (u_p, v_p) , where $u_n = u$ and $v_n = v$, and where $[(u_p - u_{p-1})^2 + (v_p - v_{p-1})^2]^{1/2} \leq \epsilon$, ($p = 1, 2, \dots, n-1$), with ϵ tending to zero as n increases without limit. Then if we put $u_p - u_{p-1} = \delta u_p$ and $v_p - v_{p-1} = \delta v_p$, we have

$$I_{j, k} = \int_{C_\lambda} d\phi_{j, k} = \lim_{n \rightarrow \infty} \sum_{p=1}^{n-1} D_{Y_{p-1} Y_p}^{(j, k)}.$$

We have therefore

THEOREM 9.1. *The integral $I_{j, k}$ is the limit of the sum of infinitesimal non-euclidean distances, each of which is defined at a separate point Y_{p-1} of C_λ as the principal part of the corresponding cross ratio $(Y_{p-1}, K, Y_p, K_{\lambda, j, k})$ which is geometrically determined at Y_{p-1} .*

This geometric characterization adds to the significance of the extremals of the special integrals which have been studied heretofore. Among these are the *pangeodesics* which are the extremals of the integral $I_{1, 1}$, and the two families of *hypergeodesics* which are the extremals of the integrals $I_{1, 0}$ and $I_{0, 1}$, respectively.

The element of projective arc length

$$(9.3) \quad ds = 2(a' b d u d v)^{1/2}$$

may be characterized geometrically in a somewhat similar manner. The line l_λ which is the R_λ -associate of l , envelops a conic C as the direction $dv/du = \lambda$ is varied while (u, v) are held constant. The conic passes through the points ρ and σ in which l intersects the u and v tangents to S at y and is in fact tangent to these asymptotic tangents at these points. The equation of the conic C when referred to the triangle of reference, whose vertices are the points y, ρ , and σ may easily be found to be

$$(9.4) \quad x_1^2 = 16a'b x_2 x_3.$$

Let Q_1, Q_2 denote the points of intersection of the tangent line to C_λ at y with the conic (9.4). If we replace, throughout the theory for the characterization of $d\phi_{j,k}$ and $I_{j,k}$, the points K and $K_{\lambda,j,k}$ by the points Q_1 and Q_2 , we obtain geometric determinations for the form $ds = 2(a'bv')^{1/2}du$ and the corresponding integral $I = \int_{C_\lambda} 2(ba'v')^{1/2}du$ known as the *projective arc-length*.

Another interesting invariant, which has been characterized by Bompiani, is $-b/a'\lambda^3 = -d\phi_{1,0}/d\phi_{0,1}$. We find that this invariant may be characterized geometrically by the cross ratio

$$(0, \infty, -bdu^2/a'dv^2, dv/du)$$

which the tangent line t_λ makes with the u and v asymptotic tangents and the R_λ -correspondent of t_λ at the point y of S .

10. Pangeodesics and union curves.* The equation of a general quadric which has contact of the second order with a surface S at a point y is

$$(10.1) \quad x_2 x_3 + x_4(-x_1 + k_2 x_2 + k_3 x_3 + k_4 x_4) = 0,$$

where the coefficients k_2, k_3, k_4 are arbitrary constants for the fixed point y and functions of u, v when y is varied over S . This quadric cuts the surface S in a curve with a triple point at y , whose tangents are in the directions satisfying the equation

$$(10.2) \quad 2bdu^3 - 3k_2 du^2 dv - 3k_3 dudv^2 + 2a'dv^3 = 0.$$

If two of these triple-point tangents coincide in the direction $dv/du = \lambda$, the third tangent must be in the direction $dv/du = -b/a'\lambda^2$. Hence we have the following

THEOREM 10.1. *If a quadric having second order contact with S at y intersects S in a curve having two coincident triple-point tangents t_λ at y , in the direction $dv/du = \lambda$, where λ is an arbitrary function of u, v , the remaining triple-point tangent is the R_λ -correspondent of t_λ .*

* Union curves were introduced by Miss P. Sperry [1, p. 214].

For each selection of λ there is a pencil of quadrics characterized by the hypothesis of the above theorem. The equation for a general one of these quadrics is given by (9.1) where

$$(10.3) \quad k_2 = (4b/\lambda - 2a'\lambda^2)/3, \quad k_3 = (4a'\lambda - 2b/\lambda^2)/3, \quad (k_4 \text{ arbitrary}).$$

Since the quadric of Moutard at y and in the direction $dv/du = \lambda$ is one of the quadrics of this pencil, we shall call this pencil the *Moutard pencil* of quadrics corresponding to the tangent t_λ at y . The following theorem, the proof of which will be left to the reader, characterizes a new line of the first kind in association with an arbitrary line of the second kind.

THEOREM 10.2. *The polar reciprocal of an arbitrary line l' of the second kind with respect to a quadric of the Moutard pencil corresponding to a tangent t_λ at y is a line $l_{m,\lambda}$ of the first kind which is dependent on the choice of l' and λ but is independent of the choice of the quadric of the pencil.*

The equations of the line $l_{m,\lambda}$ are easily found to be

$$(10.4) \quad x_1 + (\beta + [2a'\lambda^2 - 4b/\lambda]/3)x_2 + (\alpha + [2b/\lambda^2 - 4a'\lambda]/3)x_3 = 0, \quad x_4 = 0$$

where α and β are the functions associated with the point z on l' given by (1.3).

We shall call the line $l_{m,\lambda}$ the $M^{(\lambda)}$ -associate of the line l' , and the congruence $\Gamma_{m,\lambda}$ generated by $l_{m,\lambda}$ as y moves over S , the $M^{(\lambda)}$ -associate of the congruence Γ' .

DEFINITION 10.1 *A family F_λ , defined by (1.6) belongs to class \mathfrak{C}_m if the $M^{(\lambda)}$ -associate of the reciprocal of a congruence central to S is itself central to S .*

The analytic condition that F_λ belong to class \mathfrak{C}_m is, therefore, that λ satisfy the partial differential equation

$$(10.5) \quad (a'\lambda^2 - 2b/\lambda)_v = (b/\lambda^2 - 2a'\lambda)_u.$$

This is, moreover, the equation for the pangeodesics. Hence we have

THEOREM 10.3. *A curve is a pangeodesic if, and only if, it belongs to class \mathfrak{C}_m .*

Let us now obtain the second order differential equation for the curves C_λ which are characterized by the property that the $M^{(\lambda)}$ -associate of an arbitrarily chosen congruence Γ' of the second kind is the ray congruence of the associated conjugate net N_λ . Since the functions α and β associated with the ray-congruence of a conjugate net N_λ are given by

$$(10.6) \quad \alpha = ([\log \lambda]_v + 2b/\lambda^2)/2, \quad \beta = (-[\log \lambda]_u + 2a'\lambda^2)/2,$$

the conditions of the problem require that $\lambda(u, v)$ satisfy the equations

$$(10.7) \quad \begin{aligned} \beta + (2a'\lambda^3 - 4b)/3\lambda &= (-[\log \lambda]_u + 2a'\lambda^2)/2, \\ \alpha + (2b - 4a'\lambda^3)/3\lambda^2 &= ([\log \lambda]_v + 2b/\lambda^2)/2. \end{aligned}$$

If we multiply the first of these equations by -2λ and the second by $2\lambda^2$, and add corresponding sides of the resulting equations, we obtain

$$(10.8) \quad \lambda_u + \lambda\lambda_v = 2b - 2\beta\lambda + 2\alpha\lambda^2 - 2a'\lambda^3.$$

If we replace λ by dv/du and $\lambda_u + \lambda\lambda_v$ by d^2v/du^2 , we obtain

$$(10.9) \quad d^2v/du^2 = 2b - 2\beta dv/du + 2\alpha(dv/du)^2 - 2a'(dv/du)^3,$$

which is the well known differential equation for the union curves of the congruence Γ' . Hence, we have

THEOREM 10.4. *The curves defined by (1.6), which possess the property that the $M^{(n)}$ -associate of a congruence Γ' of the second kind is the ray congruence of the associated conjugate net N_λ , are union curves of the congruences Γ' .*

11. The projective normal. The purpose of this section is to present a new geometric characterization of the *projective normal*. Consider the points τ and ω , given by (3.1) and (3.2), respectively, which are the points distinct from y in which an arbitrary line l' of the second kind intersects the quadrics of Wilczynski and Lie at y . As the point y moves along a curve C_λ , the points τ and ω describe corresponding curves. The tangent lines at τ and ω to these curves intersect the tangent plane to S at y in points which we denote by T_λ and W_λ , respectively. The expression for the general coordinates of T_λ is given by a linear combination of τ and $\tau_u + \lambda\tau_v$ which does not contain y_{uv} . A similar combination of ω and $\omega_u + \lambda\omega_v$ gives the expression for the general coordinates of W_λ . The term of $\tau_u + \lambda\tau_v$ which involves y_{uv} is $-(\beta + \alpha\lambda)y_{uv}$. The same term appears in $\omega_u + \lambda\omega_v$. Hence, the expressions for T_λ and W_λ are

$$T_\lambda = \tau_u + \lambda\tau_v + (\beta + \alpha\lambda)\tau, \quad W_\lambda = \omega_u + \lambda\omega_v + (\beta + \alpha\lambda)\omega.$$

Making use of the expressions for τ and ω and the equations (1.1), we obtain

$$(11.1) \quad W_\lambda - T_\lambda = -2a'b[y_u - \tilde{\beta}y + \lambda(y_v - \tilde{\alpha}y)],$$

where $\tilde{\beta} = -\beta - (\log a'b)_u$, $\tilde{\alpha} = -\alpha - (\log a'b)_v$. Let t_λ denote the tangent to C_λ at y , let r denote the line joining W_λ and T_λ , and let ν_λ denote the point of intersection of t_λ and r . Since the right member of (11.1) is a linear combination of the expressions for the coordinates of W_λ and T_λ , it is the expression for the coordinates of ν_λ . We shall call the point ν_λ the ν -point of t_λ , associated with the line l' .

Since the right member of (11.1) is a linear combination of $y_u - \beta y$ and $y_v - \alpha y$, the point ν_λ , for any value of λ , lies on a straight line which joins \bar{p} and $\bar{\sigma}$ given by $\bar{p} = y_u - \beta y$, $\bar{\sigma} = y_v - \alpha y$, where β and α are defined above.

We shall call the line \bar{l} the ν -associate of the line l' , corresponding to the point y of S . We state now

THEOREM 11.1. *As the direction λ is varied at y , the ν -point of l_λ , associated with the line l' , describes a straight line which we call the ν -associate of the line l' .*

The point μ of intersection of the reciprocal l of l' with the line \bar{l} has general coordinates of the form

$$(11.2) \quad \mu = (\bar{\alpha} - \alpha)(y_u + [\log a'b]_u y/2) - (\bar{\beta} - \beta)(y_v + [\log a'b]_v y/2).$$

Since the intersections of the reciprocal l_n of the projective normal with the u and v tangents to S at y are given by

$$(11.3) \quad \rho = y_u + [\log a'b]_u y/2, \quad \sigma = y_v + [\log a'b]_v y/2,$$

and since μ is a linear combination of these expressions, we have

THEOREM 11.2. *The point μ , which is the intersection of the reciprocal of an arbitrarily chosen line l' of the second kind with the ν -associate of l' , lies on the reciprocal of the projective normal.*

If we consider a pair of lines l'_1, l'_2 of the second kind, we may determine the reciprocal l_n of the projective normal as the line joining the points μ_1 and μ_2 which correspond to l'_1 and l'_2 , respectively, at the point y .

Let the tangent line to S at y , which contains the point μ in correspondence with l' at y , be denoted by t . The equations for t are

$$(11.4) \quad (\bar{\alpha} - \alpha)x_3 + (\bar{\beta} - \beta)x_2 = 0, \quad x_4 = 0.$$

The equations for the lines l and \bar{l} are (1.4) and

$$(11.5) \quad x_1 - (\beta + [\log a'b]_u)x_2 - (\alpha + [\log a'b]_v)x_3 = 0, \quad x_4 = 0,$$

respectively. The harmonic conjugate of t with respect to l and \bar{l} is the line l_n whose equations are

$$(11.6) \quad 2x_1 - (\log a'b)_u x_2 - (\log a'b)_v x_3 = 0, \quad x_4 = 0.$$

Hence we have

THEOREM 11.3. *The harmonic conjugate of t with respect to the lines l and \bar{l} is the line l_n which is the reciprocal of the projective normal.*

Let l' denote the tangent to S at y which is the conjugate of t . Now since l' and the lines l, \bar{l} , and l_n which are reciprocals of l, \bar{l} , and l_n , respectively, are

also reciprocal polar lines of t , l , \bar{l} and l_n with respect to any quadric of Darboux, we have

THEOREM 11.4. *The lines t' , l' , \bar{l}' and l_n' are coplanar, and the harmonic conjugate of t' with respect to the lines l' and \bar{l}' is the projective normal.*

12. Hypergeodesics associated with the projective normal. The tangent line t which we have defined in association with an arbitrary line l' of the second kind can be used effectively to obtain new geometric characterizations for several important families of hypergeodesics associated with a surface S , namely, the projective geodesics, the union curves of the projective normal, and the dual union curves of the projective normal. These characterizations are completely described in the following theorems.

THEOREM 12.1. *The tangent t associated with the line l' which is the cuspidal axis at y of a pencil p_λ , of conjugate nets, and the tangent to the curve C_λ of the fundamental net N_λ at y are conjugate tangents if, and only if, the curve is a projective geodesic.*

According to the hypothesis we must have

$$\lambda = (2\beta + [\log a'b]_u)/(2\alpha + [\log a'b]_v),$$

where $\beta = -(\log \lambda)_u/2$ and $\alpha = (\log \lambda)_v/2$. Hence, on clearing of fractions we obtain

$$(12.1) \quad \lambda_u + \lambda\lambda_v = -(\log a'b)_v\lambda^2 + (\log a'b)_u\lambda,$$

which is the equation for the projective geodesics. The substitutions are reversible and therefore the condition is necessary and sufficient.

THEOREM 12.2. *The tangent t associated with the line l' which is the axis of y with respect to a conjugate net N_λ , and the tangent to the curve C_λ at y , are conjugate tangents if, and only if, the curve is a union curve of the projective normal.*

The hypothesis here requires that

$$\lambda = [2\beta + (\log a'b)_u]/[2\alpha + (\log a'b)_v],$$

where

$$\beta = [-(\log \lambda)_u - 2a'\lambda^2]/2, \quad \alpha = [(\log \lambda)_v - 2b/\lambda^2]/2.$$

On clearing of fractions we obtain the equation for the union curves of the projective normal

$$(12.2) \quad \lambda_u + \lambda\lambda_v = 2b + (\log a'b)_u\lambda - (\log a'b)_v\lambda^2 - 2a'\lambda^3.$$

The argument is again reversible.

The remaining theorem can be stated by replacing "the axis of y " by "the associate axis of y ," and "union curve" by "dual union curve." The method of proof is similar to that of the above theorems and will, consequently, be left to the care of the reader.

13. A system of hypergeodesics associated with an arbitrary congruence of the first kind. Consider the transversal surfaces S_{ρ_λ} and S_{σ_λ} of the R_λ -associate of an arbitrary congruence Γ of the first kind. The planes π_y and π_{ρ_λ} which are tangent at y , ρ_λ and σ_λ to the surfaces S , S_{ρ_λ} , and S_{σ_λ} , respectively, have in general the unique point in common which we denote by $P^{(\lambda)}$.

THEOREM 13.1. *If the tangent at a general point y to a curve C_λ of S contains the point $P^{(\lambda)}$ associated with an arbitrary congruence Γ at the point y , the curve is a hypergeodesic of a system which we shall call the Γ -geodesics.*

The curvilinear differential equation for the Γ -geodesics may easily be found to be

$$(13.1) \quad d^2v/du^2 = (b_u/b + \beta + \beta_{(a)})dv/du - (a'_v/a' + \alpha + \alpha_{(a)})(dv/du)^2,$$

where β , α are the functions identified with the congruence Γ , and $\beta_{(a)}$, $\alpha_{(a)}$ are those identified with the asymptotic associate of the congruence Γ .

To obtain equation (13.1) we express the condition that the tangent $yP^{(\lambda)}$ shall have the direction $dv/du = \lambda$. Since the point $P^{(\lambda)}$ is the intersection of the line joining the points ρ_λ , $(\sigma_\lambda)_{(a)}$ with the line joining the points σ_λ , $(\rho_\lambda)_{(a)}$, where $(\rho_\lambda)_{(a)}$ and $(\sigma_\lambda)_{(a)}$ determine the asymptotic associate of the line l_λ joining ρ_λ and σ_λ , the general coordinates for $P^{(\lambda)}$ may be found to be given by

$$(13.2) \quad [\alpha_\lambda - (\alpha_\lambda)_{(a)}]y_u + [\beta_\lambda - (\beta_\lambda)_{(a)}]y_v,$$

in which

$$\begin{aligned} \alpha_\lambda &= \alpha + 2a'\lambda, & \beta_\lambda &= \beta + 2b/\lambda, \\ (\alpha_\lambda)_{(a)} &= -(f + \beta_u + [2b/\lambda]_u + \beta^2 + 4b\beta/\lambda + 4b^2/\lambda^2)/2b, \\ (\beta_\lambda)_{(a)} &= -(g + \alpha_v + [2a'\lambda]_v + \alpha^2 + 4a'\alpha\lambda + 4a'^2\lambda^2)/2a'. \end{aligned}$$

The direction of the tangent $yP^{(\lambda)}$ is therefore given by

$$(13.3) \quad dv/du = [\beta_\lambda - (\beta_\lambda)_{(a)}]/[\alpha_\lambda - (\alpha_\lambda)_{(a)}].$$

On setting the right member of this equation equal to λ , simplifying and putting $\lambda = dv/du$, $\lambda_u + \lambda\lambda_v = d^2v/du^2$, we obtain equation (13.1).

The cusp-axis of the Γ -geodesics at the point y joins y to a point z given by (1.3) in which

$$\alpha = -a'_v/2a' - (\alpha + \alpha_{(a)})/2, \quad \beta = -b_u/2b - (\beta + \beta_{(a)})/2.$$

We state now two interesting theorems; the first characterizes the *directrix* of the *first kind*, of Wilczynski, and the second characterizes the *edge* of the *first kind*, of Green. The duals of these theorems characterize the corresponding lines of the second kind. The proofs of the theorems and the statements of their duals will be left to the care of the reader.

THEOREM 13.2. *The reciprocal of the cusp-axis at y of the Γ -geodesics is the directrix of the first kind, of Wilczynski, if, and only if, the reciprocal of the projective normal separates harmonically the tangent $t_{(a)}$ with respect to the lines l and $l_{(a)}$ of the congruences Γ and $\Gamma_{(a)}$, respectively.*

THEOREM 13.3. *The reciprocal of the cusp-axis of the Γ -geodesics at y , the harmonic conjugate of $t_{(a)}$ with respect to the lines l and $l_{(a)}$, and the first edge of Green, intersect in the same point, which we denote by E .*

By choosing two congruences Γ_1 and Γ_2 of the first kind, having associated with them distinct points E_1 and E_2 , we determine the first edge of Green as the line joining E_1 and E_2 .

REFERENCES

- E. ČECH
 1. *L'intorno d'un punto d'una superficie considerato dal punto di vista proiettivo*, Annali, di Matematica (3), vol. 31 (1922), pp. 191-206.
- G. FUBINI AND E. ČECH
 1. *Introduction à la Géométrie Projective Différentielle des Surfaces*, Paris, Gauthier-Villars, 1931.
- G. M. GREEN
 1. *Memoir on the general theory of surfaces and rectilinear congruences*, these Transactions, vol. 20 (1919), pp. 79-153.
- E. P. LANE
 1. *The correspondence between the tangent plane of a surface and its point of contact*, American Journal of Mathematics, vol. 48 (1926), pp. 204-214.
- P. SPERRY
 1. *Properties of a certain projectively defined two parameter family of curves on a general surface*, American Journal of Mathematics, vol. 40 (1918), pp. 213-224.
- E. J. WILCZYNSKI
 1. *Projective differential geometry of curved surfaces (second memoir)*, these Transactions, vol. 9 (1908), pp. 79-120.
 2. *Geometrical significance of isothermal conjugacy of a net of curves*, American Journal of Mathematics, vol. 42 (1920), pp. 211-221.
- UNIVERSITY OF KANSAS,
 LAWRENCE, KAN.

INVARIANCE OF THE ADMISSIBILITY OF NUMBERS UNDER CERTAIN GENERAL TYPES OF TRANSFORMATIONS*

BY

T. N. E. GREVILLE

Most physical events can be resolved, in theory at least, into a set of independent components in such a way that, when the result of each component event is known, the result of the principal event is fully determined. If the question of order does not enter into the determination so that, without changing the situation, the component events may be thought of as occurring simultaneously, the relation between the component events and the principal event can be formulated analytically in the "Verbindung" operation of von Mises. However, if the order of the component events is significant—as, for example, when the principal event is a set of tennis, the outcome depending to some extent on the order of gains and losses in a series of individual games—von Mises points out that his methods are not applicable.† It is the purpose of this paper to develop a set of transformations capable of dealing with such problems.

There will be associated with each transformation of this set a set of admissible numbers having the power of the continuum for which the property of admissibility is invariant under the given transformation. Furthermore, every denumerable subset of such transformations will be shown to have a similar property. The meaning of admissibility may be explained as follows. Assume that a one-to-one correspondence has been established between the set of all positive integers λ and the set of all sets of integers n, r_1, r_2, \dots, r_k , such that $0 < r_1 < r_2 < \dots < r_k \leq n$. Let the digits of a number u (having as digits only zeros and ones) be divided into consecutive, nonoverlapping groups of n digits each. Let T_λ denote the transformation which transforms u into a number v constituted as follows: v contains a single digit corresponding to each group of n digits of u . This digit is a one only if the r_1 th, r_2 th, \dots , r_k th digits in the corresponding group of digits of u are all ones; otherwise, it is a zero. If $p(x)$ denotes the limit of the relative frequency of ones in the number

* Presented to the Society, June 22, 1933; received by the editors March 30, 1939. This paper is a dissertation submitted in partial fulfillment of the requirements for the degree of doctor of philosophy in the University of Michigan. The writer wishes to express his appreciation to Professor A. H. Copeland for his valuable assistance in the preparation of this paper.

† von Mises [2, pp. 108–109]. References to literature are given at the end of this paper.

x , then u is said to be admissible if $p[T_\lambda(u)]$ is equal to $p^k(x)$ for every λ .*

The set of transformations to be considered herein includes as special cases all the transformations T_λ —that is, the set of operations employed by Copeland as the fundamental set for admissible numbers. By a proper choice of the fundamental set, a class of numbers having more general properties than the admissible number can be defined, and it will be shown that numbers belonging to such classes actually exist. These transformations represent certain processes followed in the classical methods of computing probabilities, and, since their properties are arrived at through rigorous mathematical developments without any assumption of "equal likelihood," they furnish a new kind of justification for the use of the customary methods in calculating the probabilities of events consisting of combinations of other events. As a by-product of more general theorems, the existence of admissible numbers having all possible probabilities is demonstrated by a new method. This also constitutes a proof of the existence of the "normale Folge" of Reichenbach, as the latter has been proved equivalent to the admissible number.

When the theory of probability is resorted to in a practical situation, it is not, as a rule, because the events under consideration are believed to be governed altogether by chance, but because there are no data, except of a statistical nature, on which to base a prediction. However, it is conceivable, in the light of modern physical researches in the quantum theory and along other lines, that the result of analyzing an event into its ultimate constituents might be a set of independent events, each governed entirely by chance. Should this prove to be the case, the admissibility of physical events would depend definitely on the invariance of admissibility under the "Verbindung" operation and under the transformations here considered, since chance events could be expected to satisfy the conditions of admissibility. The study of the relation between a given event and the set of component events of which it is made up is facilitated by certain group properties possessed by various sets of transformations which will be dealt with. The property by which the resultant of any pair (and therefore any finite number) of transformations of a given set is itself a transformation of the set will be designated as the property G . This property is possessed by the four fundamental operations of von Mises, and by the set which is the subject of this paper, as well as by a number of special subsets, including the fundamental set for admissible numbers. It is evident that the product of any finite number of sets of transformations having the property G has the property G .

* $T_\lambda(u) = [u \in (r_1 - 1, n)] \cdot [u \in (r_2 - 1, n)] \cdots [u \in (r_k - 1, n)]$. In connection with the notation, see Copeland [1, 6].

1. **The general transformation of the set R .** Let certain special permutations of zeros and ones in the digits of a number u be associated with digits 1 of a number v , and let other such permutations be associated with digits 0 of v . Such a set of permutations may be used to define a transformation on the number u , giving rise to the number v . Only permutations of finite length will be considered. Let the digits of u be divided into a sequence of mutually exclusive groups of successive digits in the following manner. If no finite number of successive digits of u , starting with the first digit, constitutes one of the special designated permutations of zeros and ones, the entire number u is considered as a single group. Otherwise, the first group consists of the smallest number of successive digits, starting with the first, which together constitute one of the specified permutations. The second group, if any, is defined in the same way, except that it begins with the digit immediately following the last digit of the preceding group, rather than with the first digit of u . Subsequent groups are similarly defined. It follows that every digit of u belongs to one and only one group, and that either the number of groups is infinite, or else there is a last group containing an infinite number of digits (provided, of course, the number u itself contains an infinite number of digits). To each group (except the last, if any) there corresponds a digit of v whose value (0 or 1) is that associated with the permutation formed by the group. If there is a last group, the number of digits of v is finite; otherwise, it is infinite. The number v obtained by this process is unique, although the converse is not true.

An illustration of a problem for which the transformations given by other writers are not adequate is the rubber of bridge, which is won by the side winning two out of three games. In this case, the "1 permutations" are 11, 101, and 011, while the "0 permutations" are 00, 010, and 100. If

$$u = 1010011101010 \dots,$$

we should divide u into groups as follows:

$$u = 101/00/11/101/010/\dots,$$

and we should have $v = 10110 \dots$.

Wald* also makes use of the notion of "0 permutations" and "1 permutations" in determining a new number whose digits depend on those of an original "collective." My procedure differs from his in two respects which are essential to the developments in this paper. First, Wald's permutations overlap, each commencing with the first term of the collective, while mine are nonoverlapping, each commencing from the end of the preceding one. Sec-

* Wald [1].

only, Wald uses the derived number solely for the purpose of making a selection (Auswahl) from the original collective, and considers only the properties of the selected sequence $u \subset v$, while I am concerned with the properties of the number v itself.

My transformation is also similar to the type employed by Reichenbach* as the fundamental set for the "normale Folge." Reichenbach's transformation is a selection which selects every digit preceded by any one of certain specified permutations. In my method, this type of selection can be approximated by a transformation defined by means of permutations consisting of the permutations associated with Reichenbach's transformation followed by a single zero or one. My transformations are more general in that they are not restricted to mere selections; his are more general in that the permutations formed from the digits of the number may overlap, and need not be consecutive.

It is obvious from the nature of the transformation that any specified permutation is superfluous which contains another specified permutation as a group of successive digits, beginning with the first. For simplicity, it will be assumed that such redundant permutations are not used. A transformation T , defined by means of two sets of specified permutations in the manner indicated, is said to belong to the set R if the set consisting of all the specified permutations (the sum of the "0" and "1" sets) satisfies the following condition, which will be called the *condition of indeterminacy*.

Any permutation whatever of zeros and ones of finite length either (i) is itself a specified permutation, or (ii) contains a specified permutation as a group of successive digits, beginning with the first, or (iii) is contained in a specified permutation as a group of successive digits, beginning with the first.

The meaning of this condition is that, no matter how many digits of u in any group have been considered without obtaining a specified permutation, there is always a possibility that by considering more digits a specified permutation will be obtained.

The set $P^{(T)}$ of specified permutations associated with the transformation T is made up of the two sets $P_1^{(T)}$ and $P_0^{(T)}$, which consist of permutations associated with the digits 1 and 0, respectively, in the number $v = T(u)$. It will be assumed that, in all the transformations discussed, neither $P_1^{(T)}$ nor $P_0^{(T)}$ is vacuous.

Associated with every transformation T of R are the probability functions

$$(1) \quad \pi^{(T)}(p) = \sum_{(h,k)} \xi_{hk}^{(T)} p^h q^k, \quad \rho^{(T)}(p) = \sum_{(h,k)} \omega_{hk}^{(T)} p^h q^k,$$

* Reichenbach [1].

in which $\xi_{hk}^{(T)}$ denotes the number of permutations of $P_1^{(T)}$ which consist of exactly h 1's and k 0's, $\omega_{hk}^{(T)}$ denotes the number of such permutations in $P_0^{(T)}$, and $\sum_{(h,k)}$ denotes summation over all nonnegative integral values of h and k . It should be noted that $\pi^{(T)}(p)$ and $\rho^{(T)}(p)$ are exactly the expressions which would be obtained by *a priori* methods for the probabilities of occurrence of the digits 1 and 0, respectively, in the number $v = T(u)$.

A transformation T of R is said to be *admissible*, or to belong to the set R_a , if

$$\pi^{(T)}(p) + \rho^{(T)}(p) \equiv 1$$

identically in p . It is problematical whether it is a sufficient condition for admissibility to have this equality hold for a set of values of p everywhere dense on the unit interval.

A transformation T of R is said to be *finite* (or to belong to the set R_f) if $P^{(T)}$ is finite.

A transformation T of R is said to be *symmetric* (or to belong to the set R_s) if the identity

$$\pi^{(T)}(p) \equiv_p \rho^{(T)}(q)$$

is satisfied. The physical interpretation of a symmetric transformation in connection with a series of games is that a player's probability of winning a rubber or set of games is related in the same way to the probability of his winning a single game, regardless of which side he takes. A sufficient but not a necessary condition that a transformation T be symmetric is that $\xi_{hk}^{(T)} = \omega_{kh}^{(T)}$ for every pair of values of h and k . The physical meaning of the latter condition would be that the rules of the rubber or set are precisely the same for both players. If T belongs to R_s ,

$$\pi^{(T)}(1/2) = \rho^{(T)}(1/2).$$

It is evident that the sets R , R_f , and R_s have the property G . The set R_a will now be shown also to possess this property.

THEOREM 1. *If T and T' are any two transformations of R_a , and T'' is a transformation such that $T''(u) = T'[T(u)]$, then T'' belongs to R_a .*

We note that

$$\begin{aligned} \pi^{(T')} [\pi^{(T)}(p)] &= \sum_{(h,k)} \xi_{hk}^{(T')} \left[\sum_{(\gamma,\delta)} \xi_{\gamma\delta}^{(T)} p^\gamma q^\delta \right]^h \left[\sum_{(\gamma,\delta)} \omega_{\gamma\delta}^{(T)} p^\gamma q^\delta \right]^k \\ &= \sum_{(h,k)} \xi_{hk}^{(T')} \left\{ \sum_{\gamma^1} \prod_{i=1}^h \xi_{\gamma_i \delta_i}^{(T)} p^{\gamma_i} q^{\delta_i} \prod_{j=1}^k \omega_{\gamma_j' \delta_j'}^{(T)} p^{\gamma_j'} q^{\delta_j'} \right\}, \end{aligned}$$

where \sum_{γ^1} denotes summation over all possible choices of the $\gamma_i, \delta_i, \gamma_i', \delta_i'$, and

δ_j' . The latter expression may be rewritten in the form

$$\sum \gamma^{\delta h k} \xi_{h k}^{(T')} \left[\prod_{i=1}^h \xi_{\gamma_i \delta_i}^{(T)} \prod_{j=1}^k \omega_{\gamma_j' \delta_j'}^{(T)} \right] p^m q^n,$$

where $m = \sum_{i=1}^h \gamma_i + \sum_{j=1}^k \gamma_j'$ and $n = \sum_{i=1}^h \delta_i + \sum_{j=1}^k \delta_j'$. Consider, for a moment, only those terms in the summation such that $m = \mu$ and $n = \nu$. If the summation is restricted to these terms,

$$\sum \gamma^{\delta h k} \xi_{h k}^{(T')} \left[\prod_{i=1}^h \xi_{\gamma_i \delta_i}^{(T)} \prod_{j=1}^k \omega_{\gamma_j' \delta_j'}^{(T)} \right]$$

is the total number of arrangements of μ 1's and ν 0's of u which will give rise to a digit 1 of $T''(u)$. Therefore, the above expression can be written

$$\sum_{(\mu, \nu)} \xi_{\mu \nu}^{(T'')} p^{\mu} q^{\nu} = \pi^{(T'')}(p).$$

Similarly, it can be shown that

$$\rho^{(T')}\left[\pi^{(T)}(p)\right] = \rho^{(T'')}(p).$$

Hence,

$$\pi^{(T'')}(p) + \rho^{(T'')}(p) = \pi^{(T')}\left[\pi^{(T)}(p)\right] + \rho^{(T')}\left[\pi^{(T)}(p)\right] \equiv_p 1,$$

since T' belongs to R_a . This proves the theorem.

2. Relation between probability and measure. The probability functions $\pi^{(T)}(p)$ and $\rho^{(T)}(p)$ are closely related to the measure of certain sets of numbers. Consider the case in which p has a rational value β/α , and let a number y in the scale of notation with radix α be so related to the binary number u that $u^{(i)} = 1$ if $y^{(i)} = 0, 1, 2, \dots, \beta - 1$, and $u^{(i)} = 0$ otherwise. The number β is not allowed to have the values 0 and α , the rational probabilities zero and unity being excluded from consideration.

THEOREM 2. *If T is any transformation of R , and a set E consists of the numbers y in the scale of α ($p = \beta/\alpha$) associated with the set of all numbers u such that the first $\mu + \nu$ digits of $v = T(u)$ exist and consist of μ 1's and ν 0's in a prescribed order, then*

$$m(E) = [\pi^{(T)}(p)]^{\mu} [\rho^{(T)}(p)]^{\nu}.$$

Let us fix our attention on a particular u such that the first $\mu + \nu$ digits of the corresponding v are in the prescribed order. Let us suppose that γ_{hk} of the first μ 1's of the digits of v result from those permutations of the digits of u which consist of h 1's and k 0's, and that δ_{hk} of the first ν 0's of v result from such permutations. Then $\sum_{(h,k)} \gamma_{hk} = \mu$ and $\sum_{(h,k)} \delta_{hk} = \nu$, and

$\sum_{(h,k)} (\gamma_{hk} + \delta_{hk})(h+k)$ digits of u (and hence of y) are required to produce these prescribed $\mu + \nu$ digits of v .

Next, let us form arbitrary decompositions of $\mu + \nu$ into sums of nonnegative integers γ_{hk} and δ_{hk} , respectively. We can assign the number pair (h, k) to any γ_{hk} of the μ 1's of the digits of v . A digit to which (h, k) is assigned is required to be produced by h 1's and k 0's of the digits of u . The total number of possible assignments is

$$\frac{\mu!}{\prod_{(h,k)} \gamma_{hk}!}.$$

The number of such assignments with respect to the ν 0's of v is

$$\frac{\nu!}{\prod_{(h,k)} \delta_{hk}!}.$$

Hence, under the above decomposition, the number of ways in which the first $\sum_{(h,k)} (\gamma_{hk} + \delta_{hk})(h+k)$ digits of y can be chosen so as to produce the prescribed $\mu + \nu$ digits of v is

$$\frac{\mu!}{\prod_{(h,k)} \gamma_{hk}!} \frac{\nu!}{\prod_{(h,k)} \delta_{hk}!} \prod_{(h,k)} [\xi_{hk}^{(T)} \beta^h (\alpha - \beta)^k]^{\gamma_{hk}} \prod_{(h,k)} [\omega_{hk}^{(T)} \beta^h (\alpha - \beta)^k]^{\delta_{hk}}.$$

It will be observed that if one of the integers $\xi_{hk}^{(T)}$ is zero and the corresponding integer γ_{hk} is not zero, then there are no ways in which the digits of y can be chosen so as to produce the prescribed $\mu + \nu$ digits of v . If, however, both $\xi_{hk}^{(T)}$ and the corresponding γ_{hk} are zero, then the ambiguous symbol 0^0 must be assigned the value 1 in order for the formula to be correct. The same is true of $\omega_{hk}^{(T)}$ and δ_{hk} .

Since the measure of the set of points corresponding to the set of numbers y for which the first $\sum_{(h,k)} (\gamma_{hk} + \delta_{hk})(h+k)$ digits are prescribed is α^{-c} , where $c = \sum_{(h,k)} (\gamma_{hk} + \delta_{hk})(h+k)$, it follows that the measure of E is

$$\begin{aligned} \sum \sum \left\{ \frac{\mu!}{\prod_{(h,k)} \gamma_{hk}!} \frac{\nu!}{\prod_{(h,k)} \delta_{hk}!} \prod_{(h,k)} [\xi_{hk}^{(T)} \beta^h (\alpha - \beta)^k]^{\gamma_{hk}} \prod_{(h,k)} [\omega_{hk}^{(T)} \beta^h (\alpha - \beta)^k]^{\delta_{hk}} \alpha^{-c} \right\} \\ = \sum \sum \left\{ \frac{\mu!}{\prod_{(h,k)} \gamma_{hk}!} \frac{\nu!}{\prod_{(h,k)} \delta_{hk}!} \prod_{(h,k)} [\xi_{hk}^{(T)} p^h q^k]^{\gamma_{hk}} \prod_{(h,k)} [\omega_{hk}^{(T)} p^h q^k]^{\delta_{hk}} \right\}, \end{aligned}$$

where the expressions in braces are to be summed for all possible decomposi-

tions of μ and ν . After application of the multinomial theorem, the above expression becomes

$$\left(\sum_{(h,k)} \xi_{hk}^{(T)} p^h q^k \right)^\mu \left(\sum_{(h,k)} \omega_{hk}^{(T)} p^h q^k \right)^\nu = [\pi^{(T)}(p)]^\mu [\rho^{(T)}(p)]^\nu = m(E).$$

The justification for this application of the multinomial theorem lies in the absolute convergence of the series for $\pi^{(T)}(p)$ and $\rho^{(T)}(p)$. To prove this, consider the case where $\mu=1$ and $\nu=0$. Under these conditions, the series for $\pi^{(T)}(p)$ would be obtained as the measure of E , without the necessity of applying the multinomial theorem. Since $m(E) \leq 1$, the series converges absolutely. A similar argument applies to the series for $\rho^{(T)}(p)$.

It follows from this theorem that any finite transformation is admissible; for, in the case of a finite transformation, it follows from the condition of indeterminacy that the first digit of v necessarily exists. Hence, for every rational p in the interval $0 < p < 1$, $\pi^{(T)}(p) + \rho^{(T)}(p) = 1$, since the left-hand member is the measure of all numbers y in the unit interval. Since, for this case, $\pi^{(T)}(p) + \rho^{(T)}(p)$ is a polynomial in p , it must be identically 1 in p .

If T belongs to R but not to R_a , there will be certain numbers y such that the corresponding number v contains only a finite number of digits, or fails entirely to exist. If T belongs to R , it follows from Theorem 2 that the measure of the set of numbers y such that the corresponding number $v = T(u)$ has at least m digits is*

$$\sum_{s=0}^m C_{m,s} [\pi^{(T)}(p)]^s [\rho^{(T)}(p)]^{m-s} = [\pi^{(T)}(p) + \rho^{(T)}(p)]^m.$$

Since the existence of the probability $p(v)$ requires that v have an infinite number of digits, the measure of the set of numbers y for which the existence of this probability is possible is

$$\lim_{m \rightarrow \infty} [\pi^{(T)}(p) + \rho^{(T)}(p)]^m,$$

which has the value 1 or 0, according as the expression within the brackets is equal to or less than unity. If T is admissible, the value is, of course, 1; otherwise, it may be either 1 or 0, as there exist transformations such that $\pi^{(T)}(p) + \rho^{(T)}(p)$ is unity for some values of p and not for others. An example of such a transformation is given at the end of the paper.

3. Admissibility of numbers obtained through transformations of the set R_a . Although the probability functions $\pi^{(T)}(p)$ and $\rho^{(T)}(p)$ were shown in the preceding section to be the measures of the sets of numbers y for which the

* $C_{m,s}$ represents the number of combinations of m things s at a time.

first digits of v are 1 and 0, respectively, the use of these functions as representing the probabilities of success and failure of the event associated with the number v will not be justified until it has been shown that $p[T(u)] = \pi^{(T)}(p)$. To show this is the purpose of the next theorem.

THEOREM 3. *For any transformation T of R_a and any rational p , ($0 < p < 1$), the set of numbers y corresponding to the set of all numbers u such that $p[T(u)] = \pi^{(T)}(p)$ has unit measure.**

Let $V = \limsup_{\mu \rightarrow \infty} |p_\mu(v) - \pi^{(T)}(p)|$, where $v = T(u)$. Then, if $E(V > \epsilon)$ denotes the set of numbers y such that $V > \epsilon$,

$$E(V > 0) = E(V > 1/2) + E(V > 1/3) + E(V > 1/4) + \dots$$

Hence the theorem will be proved if we can show that $m[E(V > \epsilon)] = 0$ for every positive number ϵ . We have

$$E(V > \epsilon) = \lim_{\mu_0 \rightarrow \infty} \sum_{\mu=\mu_0}^{\infty} E[|p_\mu(v) - \pi^{(T)}(p)| > \epsilon]$$

and hence

$$m[E(V > \epsilon)] \leq \sum_{\mu=\mu_0}^{\infty} m\{E[|p_\mu(v) - \pi^{(T)}(p)| > \epsilon]\}$$

for every positive integer μ_0 . The remainder of the proof consists in establishing the convergence of this series. From Theorem 2, it follows that

$$m\{E[|p_\mu(v) - \pi^{(T)}(p)| > \epsilon]\} = \sum_{|s/\mu - \pi^{(T)}(p)| > \epsilon} C_{\mu,s} [\pi^{(T)}(p)]^s [1 - \pi^{(T)}(p)]^{\mu-s}$$

where the expression below the summation sign indicates that the summand is to be summed for all values of s consistent with this inequality. Borel has proved† the convergence of all series of the form

$$\sum_{\mu=1}^{\infty} \sum_{|s/\mu - p| > \epsilon} C_{\mu,s} p^s q^{\mu-s},$$

where p and q are positive numbers and $p+q=1$. Hence the theorem follows.

THEOREM 4. *For any transformation T of R_a and any rational p , ($0 < p < 1$), the set of numbers y corresponding to the set of all numbers u such that u is an element of $A(p)$ ‡ and $T(u)$ is an element of $A[\pi^{(T)}(p)]$ has unit measure.*

It is evident that T_λ belongs to the set R_f —the set $P^{(T_\lambda)}$ consisting of all possible permutations of n digits; and $\pi^{(T_\lambda)}(p) = p^k$. Moreover, it follows from

* This method was used by Copeland in similar theorems. See Copeland [5, 7].

† Borel [1].

‡ $A(p)$ is the set of all admissible numbers associated with the probability p .

Theorem 1 that $T_\lambda[T(u)] = T'_\lambda(u)$, where T'_λ belongs to R_a . By Theorem 3, the set of numbers y such that $p[T'_\lambda(u)] \neq [\pi^{(T)}(p)]^k$ has zero measure. Hence, the set for which the equality $p[T'_\lambda(u)] = [\pi^{(T)}(p)]^k$ is not satisfied for every λ , that is, the set for which $T(u)$ does not belong to $A[\pi^{(T)}(p)]$ has zero measure. Since R_a includes the identity transformation T_0 , the set for which u does not belong to $A(p)$ has zero measure.

This completes the justification of formulas (1) as the probabilities associated with numbers obtained from admissible numbers by transformations of the set R_a , where the original numbers are associated with rational probabilities. It is desirable and possible to extend these results to include the case of irrational probabilities. Such an extension will be the subject of the next section.

4. Admissibility of numbers associated with irrational probabilities. The extension of the properties of the set R_a to include admissible numbers associated with all real probabilities is accomplished by the use of the property G , and by means of the following theorem.

THEOREM 5. *Corresponding to every real number θ in the interval $0 < \theta < 1$, there exists a transformation T of R_a such that $\pi^{(T)}(1/2) = \theta$.*

Let the number θ be represented as an infinite radix fraction in the scale of two.* Let the set $P^{(T)}$ consist of the permutations c_i , ($i = 1, 2, \dots$), where c_i consists of $(i-1)$ 1's followed by a single 0. The subsets $P_1^{(T)}$ and $P_0^{(T)}$ are then formed as follows. If $\theta^{(i)} = 1$, c_i belongs to $P_1^{(T)}$; otherwise, c_i belongs to $P_0^{(T)}$. For every h , $\xi_{h1}^{(T)} = \theta^{(h+1)}$, $\omega_{h1}^{(T)} = 1 - \theta^{(h+1)}$, and $\xi_{hk}^{(T)} = \omega_{hk}^{(T)} = 0$ for $k \neq 1$. Hence,

$$\pi^{(T)}(1/2) = \sum_{i=1}^{\infty} \frac{\theta^{(i)}}{2^i} = \theta.$$

Moreover, $\pi^{(T)}(p) + \rho^{(T)}(p) = \sum_{i=0}^{\infty} p^i q = 1$, for $0 < p < 1$.

It can now be shown that transformations of the set R_a give rise to admissible numbers even when applied to admissible numbers having irrational probabilities.

THEOREM 6. *Corresponding to every transformation T of R_a and every p in the interval $0 < p < 1$, there exists a nondenumerable subset E of the set $A(p)$ such that, for every number v of E , $T(v)$ belongs to $A[\pi^{(T)}(p)]$, and the corresponding set of numbers $T(v)$ is nondenumerable.*

By Theorem 5, there exists a transformation T' of R_a such that $\pi^{(T')}(1/2) = p$. Since R_a has the property G , $T[T'(u)] = T''(u)$, where T'' belongs to R_a .

* If p is expressible as a finite sum of powers of two, so that two such representations are possible, it makes no difference which is employed.

By Theorem 4, the sets of numbers y in the scale of two corresponding to the set of numbers u such that $T'(u)$ does not belong to $A(p)$, and to the set of numbers u such that $T''(u)$ does not belong to $A[\pi^{(T)}(p)]$, both have zero measure. Therefore, the set of numbers y corresponding to the set of numbers u such that $T'(u)$ belongs to $A(p)$ and $T''(u)$ belongs to $A[\pi^{(T)}(p)]$ has unit measure. Let E denote the corresponding set of numbers $v = T'(u)$. By Theorem 2, the measure of the set of numbers y corresponding to a given number $T'(u)$ is

$$\lim_{\mu \rightarrow \infty, p \rightarrow \infty} p^\mu (1 - p)^v = 0,$$

and, similarly, the measure of the set of numbers y corresponding to a given number $T''(u)$ is

$$\lim_{\mu \rightarrow \infty, p \rightarrow \infty} [\Pi^{(T)}(p)]^\mu [\rho^{(T)}(p)]^v = 0.$$

If either the set E or the corresponding set of numbers $T(v)$ were denumerable, then the set of numbers y such that v belongs to $A(p)$ and $T(v)$ belongs to $A[\pi^{(T)}(p)]$ would have the measure zero, which has been proved false.

The formulas (1), obtained originally from the *a priori* point of view, have now been justified in the light of the statistical definition of probability, since they have been shown to be the actual limiting values of the success and failure ratios associated with numbers obtained from admissible numbers by transformations of the set R_a . Theorem 6 also furnishes a new proof for the existence of admissible numbers having all probabilities in the interval $0 < p < 1$.

5. Further properties of transformations of the set R . Certain additional properties of transformations of the sets R and R_a are contained in the following theorems.

THEOREM 7. *For any transformation T of R and any p in the interval $0 < p < 1$, $\pi^{(T)}(p) + \rho^{(T)}(p) \leq 1$.*

By Theorem 5, there exists a transformation T' of R_a such that $\pi^{(T')}(1/2) = p$; and, since R has the property G , $T[T'(u)] = T''(u)$, where T'' belongs to R . Hence

$$\pi^{(T)}(p) + \rho^{(T)}(p) = \pi^{(T'')}(1/2) + \rho^{(T'')}(1/2).$$

The right-hand member has at most the value 1 since it may be regarded as the measure of a set of numbers y contained in the unit interval.

THEOREM 8. *The sets R and R_a have the power of the continuum.*

By Theorem 5, there corresponds to every number of the continuum a

distinct transformation T of R_a . It will now be shown that to every transformation of R there corresponds a distinct number of the continuum. Assume that a one-to-one correspondence has been established between the set of all positive integers λ and the set $P \equiv [c_\lambda]$ of all finite permutations of zeros and ones. Corresponding to any transformation T of R , let the number x_T be defined as follows: $x_T^{(2\lambda-1)} = 1$ if c_λ belongs to $P_1^{(T)}$, and 0 otherwise; $x_T^{(2\lambda)} = 1$ if c_λ belongs to $P_0^{(T)}$, and 0 otherwise.

6. Invariance of admissibility under transformations of the set R_a . By the definition of the admissible number, the property of admissibility is an absolute invariant under all transformations of the type T_λ . This is what is meant by the statement that the set D_a of all such transformations is the fundamental set for the set M_a of all admissible numbers. It was not, however, proved in §4 that the result of applying a transformation of the set R_a to any admissible number is always an admissible number. In fact, it is only reasonable to suppose that, if the set D_a be increased by the addition of certain other transformations of R_a , the set M_a will have to be decreased in order that the absolute invariance be preserved. It will be shown, however, that the set of numbers is not materially decreased by adding to the fundamental set any denumerable set of transformations of the set R_a . If v belongs to $A(p)$ and $T(v)$ belongs to $A[\pi^{(T)}(p)]$, the admissibility of v is said to be *regularly invariant* under the transformation T .

THEOREM 9. *If D is any denumerable subset of R_a , there exists a set M , consisting of admissible numbers and having the power of the continuum, such that, for every number of M , the property of admissibility is regularly invariant under every transformation of D .*

Let the set D consist of the transformations T_1, T_2, T_3, \dots , and let T_0 denote the identity transformation. For every p in the interval $0 < p < 1$, there exists, by Theorem 5, a transformation T of R_a such that $\pi^{(T)}(1/2) = p$. Let $N_p^{(i)}$ denote the set of numbers y in the scale of two corresponding to the set of all numbers u such that $T_i(v) = T_i[T(u)]$ does not belong to $A[\pi^{(T_i)}(p)]$. Then, by Theorem 4, $m(N_p^{(i)}) = 0$. Hence, $mC[\sum_{i=0}^{\infty} N_p^{(i)}] = 1$. Therefore, the set M_p of all numbers $v = T(u)$ associated with the numbers y of the set $C[\sum_{i=0}^{\infty} N_p^{(i)}]$ is, by the reasoning of Theorem 6, nondenumerable, and therefore nonvacuous. Let the set M consist of all the sets M_p for all probabilities p in the unit interval. It follows at once from the method of construction of M that the admissibility of every number of M is regularly invariant under every transformation of D . M has the power of the continuum since each M_p contains at least one number, which cannot belong to any other M_p .

The characteristic property of the numbers of M is not mere admissibil-

ity, but regular invariance of admissibility under the transformations of D . If thorough consistency is desired, not merely invariance of admissibility, but invariance of this characteristic property, should be demanded of the numbers of M . In other words, the result of applying a transformation of D to a number of M should be not merely an admissible number with the appropriate probability, but a number of M . In general, this type of invariance can be secured only if D has the property G . However, it is always possible to increase the set D so that it will have this property, without sacrificing denumerability. To accomplish this, add to D every transformation which is the resultant of any finite number of transformations of D . Examples of denumerable subsets of R_a having the property G are the sets R , R_f , and D_a . For all such sets the following theorem holds.

THEOREM 10. *If D is any denumerable subset of R_a having the property G , there exists a set M , consisting of admissible numbers and having the power of the continuum, such that the result of applying successively to a number of M any finite number of transformations of D is a number of M , the property of admissibility being regularly invariant under all such transformations.*

Let M be defined as in Theorem 9; and suppose there is a number u of M such that, by a finite number of transformations of D , it is possible to obtain from u a number v which does not belong to M . Since D has the property G , there is a transformation T of D , such that $v = T(u)$. Since v does not belong to M , there is a transformation T' of D such that $T'(v)$ does not belong to $A \{ \pi^{(T')} [p(v)] \}$. But

$$T'(v) = T' [T(u)] = T''(u),$$

where T'' belongs to D . This contradicts Theorem 9. The regular invariance of admissibility is an immediate consequence of the method of construction of M .

Any denumerable subset D of R_a may, therefore (with additions if necessary), play the same role in the definition of a system of numbers as the set D_a in the definition of admissible numbers. The characteristic property of such a system would be invariance of admissibility under the transformations of D . This characteristic property is itself invariant under the transformations of D , so that the system of numbers so defined is, in a sense, closed with respect to the transformations of D . It is a question whether there exists a number for which admissibility is regularly invariant under all the transformations of R_a .

7. Illustrations. (i) Let a series of games be so arranged that A wins the series if he wins a total of r games before B wins s games; otherwise B wins the series. Let u represent the sequence of games, the digit 1 denoting a game

won by A and 0 by B , let v represent the resulting sequence of rubbers or series of games, and let T denote the transformation such that $v = T(u)$. In this case, $\xi_{rk}^{(T)} = C_{r+k-1,k}$ for $k < s$, and 0 for $k \geq s$; and $\xi_{hk}^{(T)} = 0$ for $h \neq r$. Similarly, $\omega_{hs}^{(T)} = C_{h+s-1,h}$ for $h < r$ and 0 for $h \geq r$; and $\omega_{hk}^{(T)} = 0$ for $k \neq s$. Hence,

$$\pi^{(T)}(p) = \sum_{k=0}^{s-1} C_{r+k-1,k} p^r q^k, \quad \rho^{(T)}(p) = \sum_{h=0}^{r-1} C_{h+s-1,h} p^h q^s.$$

Since this transformation satisfies the condition of indeterminacy, it belongs to R_I , and therefore to R_a . An illustration of this situation is found in the game of bridge, where u represents a sequence of "games" and v the resulting sequence of rubbers. Since the rubber is won by the side first winning two games, $r = s = 2$, and

$$\pi^{(T)}(p) = p^2 + 2p^2q, \quad \rho^{(T)}(p) = q^2 + 2pq^2.$$

(ii) Let a series of games be so arranged that A wins the series if and when the number of games won by him exceeds by m the number won by B , provided the number of games won by B has not previously exceeded by n the number won by A . Similarly, B wins if he secures a lead of n games before A is m games ahead. The series is assumed to be continued until one of the players wins. Evidently this is not a finite transformation, since there is no upper limit to the number of games which may be necessary. It does, however, satisfy the condition of indeterminacy, and it can be shown* that

$$\pi^{(T)}(p) = \frac{p^m(p^n - q^n)}{p^{m+n} - q^{m+n}}, \quad \rho^{(T)}(p) = \frac{q^n(p^m - q^m)}{p^{m+n} - q^{m+n}}$$

if $p \neq q$; and if $p = q = 1/2$,

$$\pi^{(T)}(1/2) = \frac{n}{m+n}, \quad \rho^{(T)}(1/2) = \frac{m}{m+n}.$$

If $m = n$, both expressions reduce to the simpler form

$$\pi^{(T)}(p) = \frac{p^m}{p^m + q^m}, \quad \rho^{(T)}(p) = \frac{q^m}{p^m + q^m}.$$

In all cases, $\pi^{(T)}(p) + \rho^{(T)}(p) \equiv_p 1$, so that T belongs to R_a . This situation is illustrated by the case of two players matching pennies, one player starting with m and the other with n pennies, the game terminating when either player has lost all his pennies.

(iii) The following is an example of a transformation which belongs to the

* Uspensky [1, pp. 139-142].

set R but is not admissible. It also illustrates the fact that $\pi^{(T)}(p) + \rho^{(T)}(p)$ may be unity for some values of p and not for others. Let the set of specified permutations consist of all those in which the number of ones exceeds the number of zeros (after superfluous permutations have been eliminated). Each such permutation will be a "0 permutation" or "1 permutation" according as the number of 0's preceding the first 1 in the permutation is even or odd. It can be shown* that $\pi^{(T)}(p) + \rho^{(T)}(p)$ is p/q for $p \leq 1/2$, and 1 for $p \geq 1/2$.

REFERENCES

- E. BOREL
1. *Traité du Calcul des Probabilités et de ses Applications*, vol. 2, no. 1, chap. 1, Paris, 1926.
- A. H. COPELAND
1. *Admissible numbers in the theory of probability*, American Journal of Mathematics, vol. 50 (1928), pp. 535-552.
2. *Independent event histories*, American Journal of Mathematics, vol. 51 (1929), pp. 612-618.
3. *The theory of probability from the point of view of admissible numbers*, Annals of Mathematical Statistics, vol. 3 (1932), pp. 143-156.
4. *A matrix theory of measurement*, Mathematische Zeitschrift, vol. 37 (1933), pp. 542-555.
5. *Point set theory applied to the random selection of the digits of an admissible number*, American Journal of Mathematics, vol. 58 (1936), pp. 181-192.
6. *Probabilities and predictions*, Erkenntnis, vol. 6 (1937), pp. 189-203.
7. *Consistency of the conditions determining Kollektivs*, these Transactions, vol. 42 (1937), pp. 333-357.
- K. DÖRGE
1. *Zu der R. v. Mises gegebenen Begründung der Wahrscheinlichkeitsrechnung*, Mathematische Zeitschrift, vol. 32 (1930), pp. 232-258.
- E. KAMKE
1. *Über neuere Begründungen der Wahrscheinlichkeitsrechnung*, Jahresbericht der deutschen Mathematiker-Vereinigung, vol. 42 (1932), pp. 14-27.
2. *Einführung in der Wahrscheinlichkeitstheorie*, Leipzig, 1932.
- R. VON MISES
1. *Grundlagen der Wahrscheinlichkeitsrechnung*, Mathematische Zeitschrift, vol. 5 (1919), pp. 52-99.
2. *Vorlesungen aus dem Gebiete der angewandten Mathematik*, vol. 1, *Wahrscheinlichkeit*, Leipzig, 1931.
3. *Über Zahlenfolgen die ein kollektiv-ähnliches Verhalten zeigen*, Mathematische Annalen, vol. 108 (1933), pp. 757-772.
4. *Wahrscheinlichkeit, Statistik, und Wahrheit*, Vienna, 1936.
- H. REICHENBACH
1. *Axiomatik der Wahrscheinlichkeitsrechnung*, Mathematische Zeitschrift, vol. 34 (1932), pp. 568-619.
2. *Wahrscheinlichkeitslehre*, Leiden, 1935.
3. *Les fondements logiques du calcul des probabilités*, Annales de l'Institut Henri Poincaré, vol. 7 (1937), pp. 568-619.
- E. TORNIER
1. *Wahrscheinlichkeitsrechnung und Zahlentheorie*, Journal für die reine und angewandte Mathematik, vol. 160 (1929), pp. 177-198.

* Uspensky [1, p. 142].

2. *Die Axiome der Wahrscheinlichkeitsrechnung*, Journal für die reine und angewandte Mathematik, vol. 163 (1930), pp. 45-64.

J. V. USPENSKY

1. *Introduction to Mathematical Probability*, chap. 8, New York, 1937.

C. DE LA VALLÉE POUSSIN

1. *Sur l'intégrale de Lebesgue*, these Transactions, vol. 16 (1915), pp. 435-501.
2. *Intégrales de Lebesgue*, Paris, 1916.

A. WALD

1. *Die Widerspruchsfreiheit des Kollektivbegriffes der Wahrscheinlichkeitsrechnung*, Ergebnisse eines mathematischen Kolloquiums, vol. 8 (1935-1936), pp. 38-72.

UNIVERSITY OF MICHIGAN,
ANN ARBOR, MICH.

NON-COMMUTATIVE RESIDUATED LATTICES*

BY

R. P. DILWORTH

Introduction and summary. In the theory of non-commutative rings certain distinguished subrings, one-sided and two-sided ideals, play the important roles. Ideals combine under crosscut, union and multiplication and hence are an instance of a lattice over which a non-commutative multiplication is defined.† The investigation of such lattices was begun by W. Krull (Krull [3]) who discussed decomposition into isolated component ideals. Our aim in this paper differs from that of Krull in that we shall be particularly interested in the lattice structure of these domains although certain related arithmetical questions are discussed.

In Part I the properties of non-commutative multiplication and residuation over a lattice are developed. In particular it is shown that under certain general conditions each operation may be defined in terms of the other.

The second division of the paper deals with the structure of non-commutative residuated lattices in the vicinity of the unit element. It is found that this structure may be characterized to a large extent in terms of special types of distributive lattices (arithmetical and semi-arithmetical lattices). The next division contains a discussion of the arithmetical properties of non-commutative residuated lattices. In particular decompositions into primary and semi-primary elements are discussed.

Finally we investigate the case where both the ascending and descending chain conditions hold and prove some structure theorems which are analogous to the structure theorems of hypercomplex systems.

I. MULTIPLICATION AND RESIDUATION

1. Definitions and notations. The fixed lattice of elements a, b, c, \dots will be denoted by \mathfrak{S} . Sublattices will be denoted by German capitals, and Latin capitals will denote subsets of \mathfrak{S} which are not necessarily sublattices. $(,), [,], \supset$ will denote union, crosscut, and lattice division respectively. If $a \neq b$ and $a \supset x \supset b$ implies either $x = a$ or $x = b$, a is said to *cover* b and we write

* Presented to the Society in two parts: April 9, 1938, under the title *Non-commutative residuation*, and November 26, 1938, under the title *Archimedean residuated lattices*; received by the editors May 1, 1939.

† Lattices with a commutative multiplication have been investigated by Professor Morgan Ward and the author in a previous paper (Ward-Dilworth [7]).

$a > b$. If \mathfrak{S} has a unit element u , the elements covered by u are called *divisor-free* elements of \mathfrak{S} . If \mathfrak{S} has a null element it will be denoted by z .

\mathfrak{S} is said to satisfy the ascending chain condition if every chain $a_1 \subset a_2 \subset a_3 \subset \dots$ has only a finite number of distinct elements. Similarly if every descending chain $a_1 \supset a_2 \supset a_3 \supset \dots$ has only a finite number of distinct elements, \mathfrak{S} is said to satisfy the descending chain condition. \mathfrak{S} is called *archimedian* if both the ascending and descending chain conditions hold.

The direct product (Birkhoff [1]) of lattices $\mathfrak{L}_1, \mathfrak{L}_2, \dots, \mathfrak{L}_n$ is defined to be the set of vectors $a = \{a_1, a_2, \dots, a_n\}$, $a_i \in \mathfrak{L}_i$ with division defined by $a \supset b$ if and only if $a_i \supset b_i$. Union and crosscut are given by $(a, b) = \{(a_1, b_1), \dots, (a_n, b_n)\}$, $[a, b] = \{[a_1, b_1], \dots, [a_n, b_n]\}$.

2. **Multiplication.** A one-valued, binary operation xy is called a *multiplication* over \mathfrak{S} if the following postulates are satisfied:

M₁. ab lies in \mathfrak{S} whenever a and b lie in \mathfrak{S} .

M₂. $a = b$ implies $ac = bc$, $ca = cb$.

M₃. $a(b, c) = (ab, ac)$, $(a, b)c = (ac, bc)$.

M₄. $a(bc) = (ab)c$.

From M₂ and M₃ we have

(2.1) $a \supset b$ implies $ac \supset bc$ and $ca \supset cb$;

(2.2) $[ab, ac] \supset a[b, c]$, $[ac, bc] \supset [a, b]c$.

If in addition to M₁-M₄, postulate M₅ below is satisfied, \mathfrak{S} is said to be a *left ideal* lattice.

M₅. $a \supset ba$.

In a similar manner if M_{5'} is satisfied, \mathfrak{S} is said to be a *right ideal* lattice.

M_{5'}. $a \supset ab$.

If a lattice is both a left and right ideal lattice, it is called a *two-sided ideal* lattice, or simply *ideal* lattice.

Consider a lattice with unit element u over which a multiplication satisfying M₁-M₄ is defined and for which M₆ holds.

M₆. $ua = au = a$.

Then by M₃, M₅ and M_{5'} hold so that \mathfrak{S} is an ideal lattice. A lattice with unit element in which M₆ holds we call an *ideal lattice with unit*.

\mathfrak{S} is said to be commutative if it satisfies M₇.

M₇. $ab = ba$.

3. **Residuation.** Consider now an ideal lattice \mathfrak{S} in which the ascending

chain condition* holds. Let a and b be two elements of \mathfrak{S} . Then the set X of all elements $x \in \mathfrak{S}$ such that $a \supset xb$ is non-empty and closed with respect to union. Hence by the ascending chain condition X has a unit element $a \cdot b^{-1}$ which we call the *left residual* of b with respect to a . The left residual $a \cdot b^{-1}$ has the fundamental properties:

$$R_1. a \supset (a \cdot b^{-1})b.$$

$$R_2. a \supset xb \rightarrow a \cdot b^{-1} \supset x. \dagger$$

In a similar manner the *right residual* $b^{-1} \cdot a$ is defined by the following properties:

$$R_1'. a \supset b(b^{-1} \cdot a).$$

$$R_2'. a \supset bx \rightarrow b^{-1} \cdot a \supset x.$$

The two residuals are connected by the relation

$$(3.1) \quad a^{-1} \cdot (b \cdot c^{-1}) = (a^{-1} \cdot b) \cdot c^{-1}.$$

The residuals are connected with the multiplication by the formulas

$$(3.2) \quad (ab) \cdot b^{-1} \supset a, \quad a^{-1} \cdot (ab) \supset b,$$

$$(3.3) \quad a \cdot (bc)^{-1} = (a \cdot c^{-1}) \cdot b^{-1}, \quad (ab)^{-1} \cdot c = b^{-1} \cdot (a^{-1} \cdot c).$$

Some of the more important properties of the residuals are the following:

$$(3.4) \quad a \cdot (b^{-1} \cdot a)^{-1} \supset (a, b), \quad (a \cdot b^{-1})^{-1} \cdot a \supset (a, b);$$

$$(3.5) \quad [a, b] \cdot c^{-1} = [a \cdot c^{-1}, b \cdot c^{-1}], \quad a^{-1} \cdot [b, c] = [a^{-1} \cdot b, a^{-1} \cdot c];$$

$$(3.6) \quad a \cdot (b, c)^{-1} = [a \cdot b^{-1}, a \cdot c^{-1}], \quad (a, b)^{-1} \cdot c = [a^{-1} \cdot c, b^{-1} \cdot c];$$

$$(3.7) \quad (a, b) \cdot c^{-1} \supset (a \cdot c^{-1}, b \cdot c^{-1}), \quad a^{-1} \cdot (b, c) \supset (a^{-1} \cdot b, a^{-1} \cdot c);$$

$$(3.8) \quad a \supset b \rightarrow a \cdot c^{-1} \supset b \cdot c^{-1}, \quad c^{-1} \cdot a \supset c^{-1} \cdot b;$$

$$(3.9) \quad a \supset b \rightarrow c \cdot b^{-1} \supset c \cdot a^{-1}, \quad b^{-1} \cdot c \supset a^{-1} \cdot c;$$

$$(3.10) \quad a \cdot b^{-1} \supset a, \quad b^{-1} \cdot a \supset a;$$

$$(3.11) \quad a \cdot b^{-1} \supset c \Leftrightarrow c^{-1} \cdot a \supset b.$$

On the other hand, if we start with a lattice \mathfrak{S} in which the descending chain condition \ddagger holds and over which left and right residuals are defined having the properties given above, then we may define a multiplication over \mathfrak{S} satisfying M_1 - M_5 . For let a and b be two elements of \mathfrak{S} and let X be the

* This condition may be replaced by the weaker condition that every set S of elements of \mathfrak{S} have a union $u(S)$ and that $u(S)c = u(Sc)$.

\dagger The symbol \rightarrow indicates formal implication.

\ddagger As in the previous case this condition may be weakened.

set of elements x such that $x \cdot b^{-1} \supset a$. Then X is non-empty and closed with respect to crosscut, and hence by the descending chain condition has a null element ab . It can be shown that the product so defined satisfies M_1 - M_5 , and moreover is equal to the product similarly defined in terms of the right residual.

II. RESIDUATED LATTICES WITH UNIT

4. **Lattice structure.** Throughout this and the following section we shall assume that \mathfrak{S} is a lattice in which the ascending chain condition holds and having a multiplication satisfying M_1, \dots, M_6 . As a consequence of M_6 the residuals have the following properties:

$$(4.1) \quad a \supset b \Leftrightarrow a \cdot b^{-1} = u \Leftrightarrow b^{-1} \cdot a = u;$$

$$(4.2) \quad a \cdot u^{-1} = u^{-1} \cdot a = a;$$

$$(4.3) \quad (a, b) = u \rightarrow a \cdot b^{-1} = a, b^{-1} \cdot a = a.$$

Conversely, if we start with residuals having property (4.1) and define multiplication in terms of the residuals as in §3, then it is readily verified that the multiplication satisfies M_6 .

Of particular importance in the proofs that follow are the properties:

$$(4.4) \quad (b, c) = u \rightarrow (a, [b, c]) = [(a, b), (a, c)];$$

$$(4.5) \quad (b, c) = u \rightarrow ([a, b], [a, c]);$$

$$(4.6) \quad (a, b) = u, (a, c) = u \rightarrow (a, [b, c]) = u.$$

As a consequence of (4.4) and (4.6) we have the following property:

(4.7) If a_1, \dots, a_n are coprime in pairs, then

$$(c, [a_1, \dots, a_n]) = [(c, a_1), \dots, (c, a_n)].$$

Two sublattices \mathfrak{A} and \mathfrak{B} are said to be *coprime* if $a \in \mathfrak{A}$ and $b \in \mathfrak{B}$ imply $(a, b) = u$. We have then

LEMMA 4.1. Let \mathfrak{A} be the sublattice generated by the sublattices $\mathfrak{A}_1, \mathfrak{A}_2, \dots, \mathfrak{A}_n$ each of which contains u . Then \mathfrak{A} is the direct product of $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ if and only if $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ are coprime in pairs.

From the definitions of §1 it follows directly that $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ are coprime in pairs if \mathfrak{A} is the direct product of $\mathfrak{A}_1, \dots, \mathfrak{A}_n$. Let now $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ be coprime in pairs and let L denote the set of crosscuts $[a_1, \dots, a_n]$ where $a_i \in \mathfrak{A}_i$. We have clearly

$$[[a_1, \dots, a_n], [a'_1, \dots, a'_n]] = [[a_1, a'_1], [a_2, a'_2], \dots, [a_n, a'_n]].$$

Furthermore

$$([a_1, \dots, a_n], [a'_1, \dots, a'_n]) = [(a_1, [a'_1, \dots, a'_n]), \dots, (a_n, [a'_1, \dots, a'_n])] \\ = [(a_1, a'_1), \dots, (a_n, a'_n)]$$

by (4.7). Hence L is a sublattice and is thus equal to \mathfrak{A} . If $[a_1, \dots, a_n] = [a'_1, \dots, a'_n]$, then

$$a_i = (a_i, [a'_1, \dots, a'_n]) = [(a_i, a'_1), \dots, (a_i, a'_n)] = (a_i, a'_i).$$

Whence $a_i \supset a'_i$. Similarly $a'_i \supset a_i$ and hence $a_i = a'_i$. This completes the proof.

If the sublattices $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ have minimal elements, the conditions of Lemma 4.1 may be simplified.

COROLLARY. *If the sublattices $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ of Lemma 4.1 have minimal elements m_1, \dots, m_n , then \mathfrak{A} is the direct product of $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ if and only if m_1, \dots, m_n are coprime in pairs.*

From Lemma 4.1 we have immediately

LEMMA 4.2. *Any finite set of divisor-free elements generates a finite Boolean algebra.*

If there are only a finite number of divisor-free elements in \mathfrak{S} , we may speak of the Boolean algebra generated by the divisor-free elements. This is certainly the case when the descending chain condition holds in \mathfrak{S} , for we have

LEMMA 4.3. *If the descending chain condition holds in \mathfrak{S} , then there are only a finite number of divisor-free elements.*

Let $p_1, p_2, \dots, p_n, \dots$ be an infinite sequence of distinct divisor-free elements, and form the descending chain $a_1 \supset a_2 \supset a_3 \supset \dots$ where $a_i = [p_1, p_2, \dots, p_i]$. If $a_i = a_{i+1}$, then $[p_1, \dots, p_i] = [p_1, \dots, p_{i+1}]$ and hence

$$p_{i+1} = (p_{i+1}, [p_1, \dots, p_i]) = [(p_{i+1}, p_1), \dots, (p_{i+1}, p_i)] = u,$$

which is impossible. Thus $a_1 \supset a_2 \supset a_3 \supset \dots$ is an infinite descending chain.

5. We turn now to the study of the structure of a residuated lattice in the vicinity of the unit element and prove first the fundamental

THEOREM 5.1. *Let \mathfrak{S} be a residuated lattice with unit having only a finite number of divisor-free elements p_1, p_2, \dots, p_n . Moreover let \mathfrak{Q} be the direct product of chain lattices $\mathfrak{Q}_1, \dots, \mathfrak{Q}_n$ where \mathfrak{Q}_i is the chain $u \supset p_i \supset a_i \supset \dots \supset m_i$. Then if $m_k > b$ and b does not belong to \mathfrak{Q} , the sublattice generated by the elements of \mathfrak{Q} and the element b is the direct product of the chain lattices $\mathfrak{Q}_1, \dots, \mathfrak{Q}_k', \dots, \mathfrak{Q}_n$ where \mathfrak{Q}_k' is the chain lattice $u \supset p_k \supset a_k \supset \dots \supset m_k \supset b$.*

Proof. In view of the corollary to Lemma 4.1 it is sufficient to show that $(b, m_i) = u, i \neq k$. If $(b, m_i) \neq u$, there exists a divisor-free element p such that

$p \supset (b, m_i)$. Since $p \supset m_i$, we have $p = p_i$. Now $m_k \supset [m_k, p_i] \supset b$ since $p \supset b$. But $m_k \neq [m_k, p_i]$ since otherwise $p_i \supset m_k$ while $(p_i, m_k) = u$. Hence $b = [m_k, b_i]$ and b is contained in \mathfrak{L} which is contrary to assumption. Thus $(b, m_i) = u, i \neq k$.

This theorem enables us to construct certain characteristic sublattices with very simple properties. For let \mathfrak{B} be the Boolean algebra generated by the divisor-free elements of \mathfrak{S} . If a divisor-free element p of \mathfrak{B} covers an element a_1 , not contained in \mathfrak{B} , then \mathfrak{B} and a_1 generate a sublattice \mathfrak{L}_1 which is a direct product of chain lattices. If a_1 covers a_2 and a_2 does not belong to \mathfrak{L}_1 , then \mathfrak{L}_1 and a_2 generate a sublattice \mathfrak{L}_2 which is again a direct product of chain lattices. We may continue in this manner as long as we obtain elements a_i not contained in \mathfrak{L}_{i-1} . Having obtained a sublattice \mathfrak{L}_k in this manner, we may further extend it by building chains from other divisor-free elements. Thus if we call lattices which are direct products of chain lattices, *arithmetical* (Ward [5]), we see that the structure of a residuated lattice in the vicinity of the unit element is characterized to a large extent in terms of arithmetical lattices.

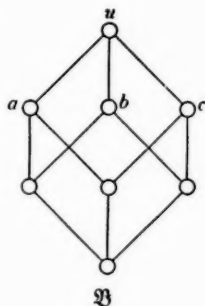


FIG. 1

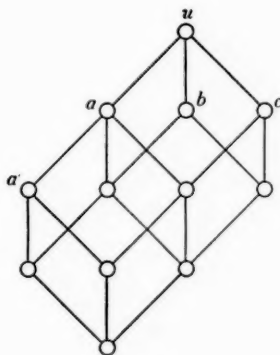


FIG. 2

This principle is very useful in constructing examples of residuated lattices. For example, suppose we wish to construct a residuated lattice containing three divisor-free elements. We start then with the Boolean algebra \mathfrak{B} of Fig. 1.

Now if we wish to add an element a' covered by a , by Theorem 5.1 we have immediately the sublattice \mathfrak{L} of Fig. 2.

The condition of Theorem 5.1 that each divisor-free element be a member of one of the chain lattices is essential for the truth of the theorem as may be seen by simple examples. However in general a residuated lattice will have

an infinite number of divisor-free elements and Theorem 5.1 will no longer apply. It may be generalized as follows:

THEOREM 5.2. *Let \mathfrak{L} be the direct product of chain lattices $\mathfrak{L}_1, \dots, \mathfrak{L}_n$ of a residuated lattice \mathfrak{S} , and let \mathfrak{B} be the lattice generated by \mathfrak{L} and the set of divisor-free elements p which divide at least one element of \mathfrak{L} . Furthermore let $m_k > b$. Then either b lies in \mathfrak{B} or the lattice generated by \mathfrak{L} and b is the direct product of the chain lattices $\mathfrak{L}_1, \dots, \mathfrak{L}'_k, \dots, \mathfrak{L}_n$ where $\mathfrak{L}'_k = \{\mathfrak{L}_k, b\}$.*

Proof. If $(b, m_i) \neq u, i \neq k$, there exists a divisor-free element p such that $p \supset (b, m_i)$. Now $m_k \supset [m_k, p] \supset b$ and $m_k \neq [m_k, p]$ since otherwise $p \supset m_k$ while $(p, m_k) = u$. Hence $b = [m_k, b]$ and $b \in \mathfrak{B}$. Hence if $b \notin \mathfrak{B}$, $(b, m_i) = u, i \neq k$, and the theorem follows by Lemma 4.

The structure of the lattice \mathfrak{B} of Theorem 5.2 is comparatively simple. We shall study its properties in terms of the notion of *semi-arithmetical* lattices introduced by Morgan Ward (Ward [5]). We make the

DEFINITION 5.1. *A distributive lattice \mathfrak{D} is said to be semi-arithmetical if the indecomposable elements divisible by a given divisor-free element form a chain lattice.*

A semi-arithmetical lattice in which the ascending chain condition holds may be characterized as follows:

LEMMA 5.1. *A distributive lattice \mathfrak{D} in which the ascending chain condition holds is semi-arithmetical if and only if the indecomposables occurring in the reduced representation of an element as a crosscut of indecomposables are coprime in pairs.*

From Definition 5.1 it follows trivially that an arithmetical lattice is semi-arithmetical.

We shall show now that the lattice \mathfrak{B} of Theorem 5.2 is semi-arithmetical and to that end prove the

THEOREM 5.3. *Let \mathfrak{L} be a semi-arithmetical sublattice of a residuated lattice \mathfrak{S} and let \mathfrak{L} contain the unit element u . Then if p is a divisor-free element of \mathfrak{S} , the sublattice \mathfrak{L}' generated by p and the sublattice \mathfrak{L} is semi-arithmetical.*

Proof. If p is contained in \mathfrak{L} , the theorem is trivial and we may thus assume that $p \notin \mathfrak{L}$. Now let U be the set of all elements of the form a or $[p, a]$ where $a \in \mathfrak{L}$. The set U is clearly closed with respect to crosscut. We show that U is also closed with respect to union. Let x and y be two members of the set U . If both x and y are contained in \mathfrak{L} , (x, y) is obviously in U . Let $x = [p, x_1]$, $p \nmid x_1$ and $y \in \mathfrak{L}$. Let $x_1 = [q_1, \dots, q_s]$ where the q_i are indecomposables and $(q_i, q_j) = u, i \neq j$. Then since $p \nmid x_1$, $(p, q_i) = u$ ($i = 1, \dots, s$). Hence

$$(x, y) = (y, [p, q_1, \dots, q_s]) = [(y, p), (y, q_1), \dots, (y, q_s)]$$

by (4.3). But (y, p) is either p or u hence (x, y) is contained in U . If $x = [p, x_1]$, $p \nmid x_1$ and $y = [p, y_1]$, $p \nmid y_1$, then

$$\begin{aligned}(x, y) &= ([p, q_1, \dots, q_s], [p, q'_1, \dots, q'_s]) \\ &= [p, (q_1, p), \dots, (q_s, p), \dots, (q'_s, p), (q_1, q'_1), \dots, (q_s, q'_s)] = [p, a]\end{aligned}$$

where $a \in \mathfrak{Q}$. Hence U is identical with \mathfrak{Q}' .

Now let a, b, c be contained in U . Then in exactly the same manner as above we find that $(a, [b, c]) = [(a, b), (a, c)]$. For example, if $b = [p, b_1]$, $p \nmid b_1$ and $c \in \mathfrak{Q}$, then

$$\begin{aligned}(a, [b, c]) &= (a, [p, q_1, \dots, q_s, q'_1, \dots, q'_s]) \\ &= [(a, p), (a, q_1), \dots, (a, q'_s)] = [(a, b), (a, c)]\end{aligned}$$

if $p \nmid c$; and if $p \mid c$, then

$$\begin{aligned}(a, [b, c]) &= (a, [q_1, \dots, q_s, q'_1, \dots, q'_s]) \\ &= [(a, q_1), \dots, (a, q_s), (a, q'_1), \dots, (a, q'_s)] \\ &= [(a, q_1), \dots, (a, q_s), (a, c)] = [(a, p), (a, q_1), \dots, (a, q_s), (a, c)] \\ &= [(a, b), (a, c)].\end{aligned}$$

Hence \mathfrak{Q}' is distributive.

Finally let $x \in \mathfrak{Q}'$; then either $x \in \mathfrak{Q}$ or $x = [p, x_1]$ where $p \nmid x_1$. If $x \in \mathfrak{Q}$, then $x = [q_1, \dots, q_r]$ where the q_i are indecomposable and $(q_i, q_j) = u$, $i \neq j$. If $x = [p, x_1]$ then $x = [p, q_1, \dots, q_r]$ where p, q_1, \dots, q_r are indecomposable and $(q_i, q_j) = u$, $i \neq j$; $(p, q_i) = u$ ($i = 1, \dots, r$). Thus \mathfrak{Q}' is semi-arithmetical by Lemma 5.1 and the proof is complete.

Now since \mathfrak{B} is obtained from an arithmetical lattice \mathfrak{Q} by a successive adjunction of divisor-free elements and since at each stage a semi-arithmetical sublattice is obtained, \mathfrak{B} itself is semi-arithmetical. We have thus proved

THEOREM 5.4. *The lattice \mathfrak{B} of Theorem 5.2 is a semi-arithmetical sublattice of \mathfrak{S} .*

In forming the sublattice \mathfrak{B} from the arithmetical lattice \mathfrak{Q} only divisor-free elements which are divisors of some element of \mathfrak{Q} are considered. If we adjoin a divisor-free element which does not divide any of the elements of \mathfrak{Q} , the results are even simpler; for we have

THEOREM 5.5. *Let \mathfrak{Q} be a direct product of the chain lattices $\mathfrak{Q}_1, \dots, \mathfrak{Q}_n$, and let p denote a divisor-free element not contained in \mathfrak{Q} . Then if p does not divide any of the elements of \mathfrak{Q} , the sublattice generated by p and \mathfrak{Q} is the direct product \mathfrak{Q}'*

of the chain $\{u, p\}$ and the chain lattices of \mathfrak{L} . Furthermore if \mathfrak{L} is dense in \mathfrak{S} , then \mathfrak{L}' is dense in \mathfrak{S} .

Proof. Since p does not divide a_i if $a_i \in \mathfrak{L}_i$, $(a_i, p) = u$. Hence the first part of the theorem follows. Let now $x \supset [p, a_1, \dots, a_n]$. Then $x = [(x, p), (x, a_1), \dots, (x, a_n)]$. Now (x, p) is clearly in \mathfrak{L}' and (x, a_i) is in \mathfrak{L} by hypothesis. Hence $x \in \mathfrak{L}'$.

We conclude this section with

THEOREM 5.6. Let \mathfrak{L} be the direct product of the chain lattices $\mathfrak{L}_1, \dots, \mathfrak{L}_n$ of a residuated lattice \mathfrak{S} and let $m_k > b$ where b is indecomposable. Then \mathfrak{L} and b generate a sublattice \mathfrak{L}' which is the direct product of the chain lattices $\mathfrak{L}_1, \dots, \{\mathfrak{L}_k, b\}, \dots, \mathfrak{L}_n$. Furthermore if \mathfrak{L} is dense in \mathfrak{S} , then \mathfrak{L}' is dense in \mathfrak{S} .

Proof. The first part follows directly from Theorem 5.2. Let now $x \supset [m_1, m_2, \dots, b, \dots, m_n]$. Then $x = [(x, m_1), \dots, (x, b), \dots, (x, m_n)]$. Since \mathfrak{L} is dense by hypothesis, $(x, m_1), \dots, (x, m_n)$ are contained in \mathfrak{L} . Now either $(x, b) = b$ in which case $x \in \mathfrak{L}'$ or $(x, b) \supset m_i$ since b is indecomposable. But then $(x, b) \in \mathfrak{L}$ and x is contained in \mathfrak{L}' .

III. ARITHMETICAL PROPERTIES OF IDEAL LATTICES

6. Assume that \mathfrak{S} is an ideal lattice in which the ascending chain condition holds.

DEFINITION 6.1. An element $p \in \mathfrak{S}$ is said to be a prime if $p \supset ab$ and $p \nmid a$ implies $p \supset b$.

DEFINITION 6.2. An element $q \in \mathfrak{S}$ is said to be right primary if $q \supset ab$ and $q \nmid a$ implies $q \supset b^s$ for some whole number s .

In the theory of commutative residuated lattices a residuated lattice in which the ascending chain condition holds is said to be a Noether lattice (Ward-Dilworth [7]) if every irreducible is primary. It is then shown that every element of a Noether lattice may be represented as a simple* crosscut of a finite number of primaries each of which is associated with a different prime. The primes themselves and the total number of primaries are uniquely determined by the element. This result also holds for the non-commutative case although there are certain complications due to the non-commutativity of the multiplication. We shall show how these complications may be avoided.

Let \mathfrak{S} be a non-commutative Noether lattice; that is, assume that every irreducible is right primary. If a and b are elements of \mathfrak{S} , the product ab then has the form $ab = [q_1, \dots, q_r]$ where the q_i are right primary. Let $q_i \supset a$

* A crosscut representation is said to be simple if omitting any one of the terms changes the representation.

($i=1, \dots, l$), $q_i \nmid a$ ($i=l+1, \dots, r$). Then since $q_i \supset ab$ we have $q_i \supset b^{s_i}$ ($i=l+1, \dots, r$). If we then set $s = \max(s_{l+1}, \dots, s_r)$, we have

$$(6.1) \quad ab \supset [a, b^s] \supset b^s a.$$

Let q be right primary and consider the union p of all elements x such that $q \supset x^s$ for some whole number s . Then $q \supset p^t$ for some whole number t by the ascending chain condition. Furthermore p is a prime. For if $p \supset ab$, then $q \supset p^t \supset (ab)^t \supset a^t b^t$ by (6.1). If $q \supset a^t$, then $p \supset a$. If $q \nmid a^t$, then $q \supset b^{ts}$ and $p \supset b$. Hence either $p \supset a$ or $p \supset b$. This prime is clearly unique and is called the prime element associated with the right primary q . We have moreover

LEMMA 6.1. *The crosscut of two right primaries associated with the same prime p is also a right primary associated with p .*

Let $[q, q'] \supset ab$, $[q, q'] \nmid a$. Then either q or q' , say q , does not divide a and hence $q \supset b^s$. But then $p \supset b$ and hence $q' \supset b^t$. Hence $[q, q'] \supset b^{st}$ where $s' = \max(s, t)$. Obviously $[q, q']$ is associated with p .

LEMMA 6.2. *Let q and q' be right primaries associated with p and p' respectively. Then if $p \nmid p'$, $q \cdot q'^{-1} = q$.*

For $q \supset (q \cdot q'^{-1})q'$. Hence either $q = q \cdot q'^{-1}$ or $q \supset q'^s$. But if $q \supset q'^s$, then $p \supset p'^s$ and hence $p \supset p'$ contrary to hypothesis.

Note that Lemma 6.2 holds only for the right residual. If we were considering left primaries, the left residual would replace the right residual.

The proof from this point on is exactly analogous to the proof in classical ideal theory and will be omitted. We thus obtain

THEOREM 6.1. *Let \mathfrak{S} be a non-commutative Noether lattice. Then every element of \mathfrak{S} may be represented as a simple crosscut of a finite number of right primaries. The primes and the total number of right primaries are uniquely determined by the element.*

The following theorem proved in Ward-Dilworth [7] for the commutative case holds also for non-commutative residuated lattices and is proved in exactly the same manner.

THEOREM 6.2. *The following two conditions are sufficient that \mathfrak{S} be a Noether lattice:*

- (i) \mathfrak{S} is modular,
- (ii) $ab \supset [a, b^s]$.

The distinction between left and right primaries may be removed by weakening the condition of Definition 6.2. We adopt the name *semi-primaries* for these new elements.

DEFINITION 6.3. An element $a \in \mathcal{S}$ is said to be semi-primary if $a \supset bc$ and $a \not\supset b^s$ for all s implies $a \supset c^t$ for some whole number t .

Let \mathcal{S} be an ideal lattice in which every element may be represented as a crosscut of a finite number of semi-primaries. Moreover let x and y be any two elements of \mathcal{S} . Then $xy = [a_1, \dots, a_r]$ where the a_i are semi-primary. Let $a_i \supset x^{s_i}$ for $i = 1, \dots, l$ and $a_i \supset y^{t_i}$, $i = l+1, \dots, r$. Then $xy \supset [x^s, y^t]$ where $s = \max(s_1, \dots, s_l)$ and $t = \max(t_{l+1}, \dots, t_r)$. We thus have

THEOREM 6.3. If every element of a residuated lattice \mathcal{S} is expressible as a crosscut of a finite number of semi-primaries, then for every x and y in \mathcal{S} , there exist whole numbers s and t such that

$$(6.2) \quad xy \supset [x^s, y^t].$$

If (6.2) holds in a residuated lattice, the semi-primary elements may be simply characterized as follows:

THEOREM 6.4. Let \mathcal{S} be a residuated lattice in which (6.2) holds. Then an element a is semi-primary if and only if a prime p exists such that $p \supset a \supset p^s$ for some whole number s .

Proof. Let a be semi-primary, and let p denote the union of all elements x such that $a \supset x^r$ for some r . Then $a \supset p^t$ for some t . Now let $p \supset xy$. Then $a \supset xy \supset x^m y^n$ for some integers m and n by (6.2). Hence $a \supset x^s$ for some s or $a \supset y^t$ for some t . Hence either $p \supset x$ or $p \supset y$. Clearly $p \supset a \supset p^s$ for some s .

Conversely let $p \supset a \supset p^s$ and suppose that $a \supset bc$. Then $p \supset bc$, and hence either $p \supset a$ or $p \supset b$. Hence either $a \supset b^s$ or $a \supset c^s$.

The converse to Theorem 6.3 does not hold in general. However under the assumption of the distributive law we have

THEOREM 6.5. The following two conditions are sufficient that every element of a residuated lattice \mathcal{S} satisfying the ascending chain condition be expressible as a crosscut of a finite number of semi-primaries.

- (i) \mathcal{S} is distributive,
- (ii) $xy \supset [x^s, y^t]$ for suitable s and t .

Every element of \mathcal{S} is clearly expressible as a crosscut of a finite number of indecomposables. Hence it is sufficient to show that every indecomposable is semi-primary. Let a be indecomposable, and let $a \supset bc$, $a \not\supset b^s$, for any s . Then $a \supset [b^s, c^t]$ by (ii). Hence $a = [(a, b^s), (a, c^t)]$ by (i). But $(a, b^s) \neq a$. Hence since a is indecomposable, $a = (a, c^t)$ and $a \supset c^t$.

The distributive condition is essential in Theorem 6.5 as is shown by the example in Fig. 3.

Let \mathfrak{L}_a denote the sublattice $\{a', b'', a, c'', b''', z', c''', d', e', z\}$, \mathfrak{L}_b the sublattice $\{d, b', b\}$, and \mathfrak{L}_c the sublattice $\{e, c', c\}$. We define a multiplication over \mathfrak{L} as follows: $u^2 = u$, $ux = b$ if $x \in \mathfrak{L}_b$, $ux = c$ if $x \in \mathfrak{L}_c$, $ux = z$ if $x \in \mathfrak{L}_a$. The product of any two elements in \mathfrak{L}_b is b . The product of any two elements in \mathfrak{L}_c is c . The product of any element of \mathfrak{L}_b with an element of \mathfrak{L}_c is z . The product of an element of \mathfrak{L} with an element of \mathfrak{L}_c is z . It is readily verified that the multiplication so defined satisfies M_1, \dots, M_8 , and is also commutative. \mathfrak{L} is

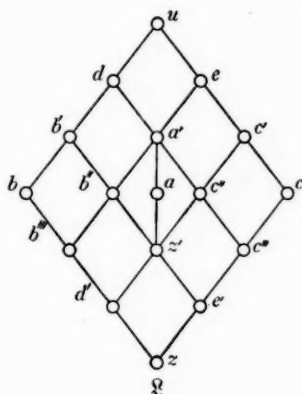


FIG. 3

clearly *not* distributive. It can also be verified that $xy \supset [x^s, y^t]$ for suitable s and t . However it is not true that $xy \supset [x, y^s]$ for some s , since $dc \not\supset [d, c^s]$. Furthermore a is indecomposable but *not* semi-primary since $a \supset bc$, but $a \not\supset b^s$ any s and $a \not\supset c^t$ any t .

7. Ideal lattices with unit. We turn now to the study of the properties of divisor-free elements in an ideal lattice with unit. We prove first the

LEMMA 7.1. *Let f be a divisor-free element of \mathfrak{L} , and let a be any element not divisible by f . Then one and only one of the following formulas holds:*

- (1) $fa \supset af$,
- (2) $fa = (fa) \cdot f^{-1}$.

We have $(fa \cdot f^{-1})^{-1} \cdot fa \supset f$ by (4.4). Hence either $(fa \cdot f^{-1})^{-1} \cdot fa = u$ or $(fa \cdot f^{-1})^{-1} \cdot fa = f$. In the first case $fa \supset fa \cdot f^{-1}$. But $fa \cdot f^{-1} \supset fa$ by (3.10). Hence $fa = fa \cdot f^{-1}$. If $(fa \cdot f^{-1})^{-1} \cdot fa = f$, then

$$f = f \cdot a^{-1} = ((fa \cdot f^{-1})^{-1} \cdot fa) \cdot a^{-1} = (fa \cdot f^{-1})^{-1} \cdot (fa \cdot a^{-1}) \supset (fa \cdot f^{-1})^{-1} \cdot f.$$

But $(fa \cdot f^{-1})^{-1} \cdot f \supset f$. Hence $(fa \cdot f^{-1})^{-1} \cdot f = f$. But then $f^{-1} \cdot (fa \cdot f^{-1}) = fa \cdot f^{-1}$ or $(f^{-1} \cdot fa) \cdot f^{-1} = fa \cdot f^{-1}$. Then $fa \cdot f^{-1} \supset a \cdot f^{-1} \supset a$. Hence $fa \supset (fa \cdot f^{-1})f \supset af$.

If both (1) and (2) hold, then $fa = fa \cdot f^{-1} \supset af \cdot f^{-1} \supset a$. But then $f = (f, fa) \supset (f, a) = u$, contrary to the assumption that f is a divisor-free element.

We clearly have a similar result for left residuals.

LEMMA 7.2. *Let f be a divisor-free element of a residuated lattice in which (6.2) holds. Then f commutes with every element which it does not divide.*

Let $a \in \mathcal{S}$ such that $f \nmid a$. Then by Lemma 7.1 either $fa \supset af$ or $fa = fa \cdot f^{-1}$. If $fa = fa \cdot f^{-1}$, then $fa \supset (a^i f^i) \cdot f^{-1} \supset a^i f^{i-1}$ by (6.2). But then $fa \supset fa \cdot f^{-1} \supset (a^i f^{i-1}) \cdot f^{-1} \supset a^i f^{i-2}$. Continuing in this manner we finally get $fa \supset a^i$. But then $f = (f, fa) \supset (f, a^i) = u$ since $f \nmid a^i$. Hence $f = u$ which is contrary to our assumption that f is a divisor-free element. We thus have $fa \supset af$. In a similar manner using left residuals we get $af \supset fa$. Hence $af = fa$.

As a corollary to Lemma 7.2 the divisor-free elements in a residuated lattice for which (6.2) holds always commute. In particular we have from Theorem 6.3

LEMMA 7.3. *If in a residuated lattice every element is expressible as a crosscut of semi-primaries, then the divisor-free elements commute.*

Let \mathcal{S} be an arbitrary residuated lattice in which the ascending chain condition holds and denote by \mathcal{S}' the set of all elements x which divide a finite product of divisor-free elements. \mathcal{S}' is clearly closed under union, crosscut, multiplication and residuation and hence a residuated sublattice of \mathcal{S} . Then

LEMMA 7.4. *Every prime in \mathcal{S}' is divisor-free.*

Let p be a prime in \mathcal{S}' . Then by the definition of \mathcal{S}' , $p \supset f_1 f_2 \cdots f_r$, where f_1, f_2, \dots, f_r are divisor-free elements of \mathcal{S} . Hence $p = f_i$ for some i .

LEMMA 7.5. *Every element of \mathcal{S}' divides a finite product of its divisor-free divisors.*

This lemma follows directly from the following lemma due to Krull [3].

LEMMA 7.6. *Let \mathcal{S} be a non-commutative residuated lattice in which the ascending chain condition holds. Then each element $a \in \mathcal{S}$ has only a finite number of minimal prime divisors p_1, \dots, p_n and a divides a power of $p_1 \cdots p_n$.**

* Krull states this lemma for the more general case where the ascending chain condition is assumed only for prime elements while a residual chain condition holds for all elements. However his proof seems to be in error as he uses the following rule: If $a \supset a_1' a_2'$, then $a \supset a_1 a_2$ where $a_1 = a \cdot a_2'^{-1}$ and $a_2 = a_1'^{-1} \cdot a$. This rule is in general not correct as the following example shows: Let \mathcal{S} be the lattice defined by the covering relations $u > a > b > c > z$, $b > d > z$. The multiplication is defined by $ux = xu = x$, all $x \in \mathcal{S}$; $a^2 = a$, and all other products are equal to z . Then $z \cdot c^{-1} = a$, $d^{-1} \cdot z = a$ and $z \supset cd$. However $z \nmid (z \cdot c^{-1})(d^{-1} \cdot z) = a^2 = a$.

The lemma is readily seen to be correct under the assumption of the ascending chain condition since we may take $a_1 = (a, a_1')$ and $a_2 = (a, a_2')$ and the rule stated above holds.

A further consequence of Lemma 7.6 is the result that \mathfrak{S}' is the maximal residuated sublattice all of whose prime elements are divisor-free.

In certain cases \mathfrak{S}' is simply the Boolean algebra \mathfrak{B} generated by the divisor-free elements. For example we have

THEOREM 7.1. *Let \mathfrak{S} be a residuated lattice with only a finite number of divisor-free elements all of which commute among themselves. If the only elements covered by the divisor-free elements are elements of the Boolean algebra \mathfrak{B} generated by them, then $\mathfrak{S}' = \mathfrak{B}$.*

Proof. Under the hypothesis of the theorem, $f^2 \leq [f, f']$ or $f^2 = f$. But if $[f, f'] \geq f^2$, then $f' \geq f$ which is impossible. Hence $f^2 = f$. But then $[f_1, f_2, \dots, f_n] = f_1 f_2 \dots f_n$ and $(f_1 \dots f_n)^2 = f_1 \dots f_n$.

If the divisor-free elements do not commute, the theorem does not hold in general. Consider the lattice \mathfrak{L} defined by the covering relations $u > b > c > z$, $u > a > c$. The multiplication is given by $ux = xu = x$, $x \in \mathfrak{L}$, and $ab = c$, $ba = z$, $ac = ca = bc = cb = c^2 = z$, $a^2 = a$, $b^2 = b$, $zx = z$, all $x \in \mathfrak{L}$. Then $\mathfrak{S}' = \mathfrak{L}$ while \mathfrak{B} is the sublattice $\{u, a, b, c\}$.

Applying Theorem 7.1 to hypercomplex systems we obtain

THEOREM 7.2. *A hypercomplex system in which the prime two-sided ideals are commutative is a direct sum of simple two-sided ideals if and only if each irreducible two-sided ideal which is not a prime has at least two prime ideal divisors.*

We conclude this section by giving a variation of a theorem due to Krull.*

THEOREM 7.3. *Each element of \mathfrak{S}' is expressible as a crosscut of a finite number of semi-primaries if and only if the divisor-free elements commute.*

Proof. The second part follows from Lemma 7.3. To prove the first let $a = [a_1, \dots, a_r]$ be the decomposition of a into coprime indecomposable elements. Then $a_i \geq f_1^{n_1} \dots f_r^{n_r} = [f_1^{n_1}, \dots, f_r^{n_r}]$ or $a_i = [(a_i, f_1^{n_1}), \dots, (a_i, f_r^{n_r})]$ whence $a_i = (a_i, f_j^{n_j})$ for some j . We have then $f_j \geq a_i \geq f_j^{n_j}$. Let $a_i \geq bc$; then $f_j \geq bc$ and hence either $f_j \geq b$ or $f_j \geq c$. Hence either $a_i \geq b^{n_j}$ or $a_i \geq c^{n_j}$. Thus if the divisor-free elements of \mathfrak{S} commute, each element of \mathfrak{S}' may be uniquely represented as a crosscut of coprime semi-primary elements.

IV. ARCHIMEDEAN RESIDUATED LATTICES

8. Throughout this section unless the contrary is explicitly stated it will be assumed that \mathfrak{S} is an ideal lattice in which the ascending and descending chain conditions hold. The unit element of \mathfrak{S} need not be the unit of multiplication.

* Krull proves the theorem for "primary" elements where an element is primary if it has only one divisor-free divisor.

DEFINITION 8.1. An element a of \mathfrak{S} is said to be nilpotent if $a^s = z$ for some whole number s .

LEMMA 8.1. The union m of all nilpotent elements of \mathfrak{S} is nilpotent. m is called the radical of \mathfrak{S} .

If $a_1^{t_1} = z$ and $a_2^{t_2} = z$, then $(a_1, a_2)^t = z$ where $t = t_1 + t_2 - 1$. The result follows from the ascending chain condition.

DEFINITION 8.2. An element s of \mathfrak{S} is said to be simple if $s > z$ where z is the null element of \mathfrak{S} .

LEMMA 8.2. A necessary and sufficient condition that the radical be the null element is that each simple element be idempotent.

Let $m = z$. If s is a simple element, since $s \supset s^2$, either $s = s^2$ or $s^2 = z$. But if $s^2 = z$, then $z \supset m \supset s$ contrary to Definition 8.2. Suppose now that each simple element is idempotent and let $m \neq z$. Then $m \supset s$ where s is simple, whence $z = m' \supset s' = s$, which contradicts the definition of s . Hence $m = z$.

DEFINITION 8.3. If the radical is the null element, \mathfrak{S} is said to be semisimple.

LEMMA 8.3. Let \mathfrak{S} be semisimple and s be any simple element of \mathfrak{S} . Then $a \supset s \Leftrightarrow as = sa = s$, $a \not\supset s \Leftrightarrow as = sa = z$.

Let $a \supset s$. Then $as \supset s^2 = s$ and hence $as = s$. Similarly $sa = s$. If $a \not\supset s$, then $[a, s] = z$ and hence $as = sa = z$.

The position of the radical in the lattice may have important bearing on the arithmetical properties of the lattice. For example, we have the following theorem:

THEOREM 8.1. Let \mathfrak{S} be an archimedean residuated lattice whose divisor-free elements generate a Boolean algebra with null element m . Then the divisor-free elements are the only primes of \mathfrak{S} .

Proof. Since \mathfrak{S} is archimedean there is only a finite number of divisor-free elements. Let p be a prime of \mathfrak{S} . Then $p \supset m' \supset z$ and hence $p \supset m$. But $m = [f_1, \dots, f_n]$ where f_1, \dots, f_n are the divisor-free elements of \mathfrak{S} . Hence $p \supset [f_1, \dots, f_n]$ and hence $p = f_i$ for some i .

The conclusion of Theorem 8.1 may be stated in the form $\mathfrak{S} = \mathfrak{S}'$.

Let \mathfrak{S}_m denote the sublattice of all elements x such that $x \supset m$. The study of the structure of \mathfrak{S}_m may be reduced to the study of the structure of semisimple lattices. For since \mathfrak{S}_m is dense in \mathfrak{S} it is closed with respect to residuation and hence has a multiplication (§3). We call this multiplication the multiplication in \mathfrak{S}_m and denote it by $a \cdot b$.

THEOREM 8.2. Let $a, b \in \mathfrak{S}_m$. If $ab \in \mathfrak{S}_m$, then $ab = a \cdot b$.

Proof. $a \cdot b$ is defined by

- (i) $(a \cdot b) \cdot b^{-1} \supset a$,
- (ii) $x \cdot b^{-1} \supset a, x \in \mathfrak{S}_m \rightarrow x \supset a \cdot b$.

Similarly ab is defined by

- (i') $(ab) \cdot b^{-1} \supset a$,
- (ii') $x \cdot b^{-1} \supset a, x \in \mathfrak{S} \rightarrow x \supset ab$.

Hence if $ab \in \mathfrak{S}_m$, then $ab \supset a \cdot b$ by (i'), (ii). On the other hand by (i), (ii'), $a \cdot b \supset ab$. Hence $a \cdot b = ab$.

In general we have

LEMMA 8.4. $a \cdot b \supset ab$.

Let now p be a prime element of \mathfrak{S} . Then $p \supset m' = z$ and hence $p \supset m$. Thus $p \in \mathfrak{S}_m$. Now let $p \supset a \cdot b$. Then $p \supset ab$ by Lemma 8.4. Hence either $p \supset a$ or $p \supset b$. We thus have

THEOREM 8.3. *If p is a prime element of \mathfrak{S} , then $p \in \mathfrak{S}_m$ and p is a prime in \mathfrak{S}_m with respect to the multiplication in \mathfrak{S}_m .*

THEOREM 8.4. \mathfrak{S}_m is semisimple.

Proof. Let s be a simple element of \mathfrak{S}_m . Then $s \supset m$. Now $s \supset s \cdot s$. Hence $s = s \cdot s$ or $s \cdot s = m$. But if $s \cdot s = m$, $m \supset s^2$ by Lemma 8.4 and hence $s^2 = z$. This contradicts the definition of m . Hence each simple element is idempotent and by Lemma 8.1 \mathfrak{S}_m is semisimple.

The most important application of archimedean residuated lattices is in the theory of hypercomplex systems. More generally, let \mathfrak{S} be the set of two-sided ideals of a non-commutative ring R in which the ascending and descending chain conditions hold for left ideals. Then m is the radical of R . Now the quotient ring R/m is isomorphic to \mathfrak{S}_m and hence is semisimple by Theorem 8.4. However from a well known structure theorem, a semisimple ring is a direct sum of simple two-sided ideals. Its lattice of two-sided ideals is thus a Boolean algebra, and Theorem 8.1 gives

THEOREM 8.5. *The only prime two-sided ideals in a hypercomplex system are the divisor-free ideals.*

9. Semisimple lattices. In this section we shall be particularly interested in the sublattices generated by the simple elements of a semisimple lattice \mathfrak{S} .

LEMMA 9.1. *There are only a finite number of simple elements in a semisimple lattice \mathfrak{S} .*

Let s_1, s_2, s_3, \dots be an infinite sequence of simple elements. Consider the chain $a_1 \subset a_2 \subset a_3 \subset \dots$ where $a_i = (s_1, s_2, \dots, s_i)$. The members of this chain are distinct. For suppose that $a_i = a_{i+1}$; then $(s_1, \dots, s_i) = (s_1, \dots, s_{i+1})$.

Hence we have

$$\begin{aligned} s_{i+1} &= s_{i+1}^2 = (s_1 s_{i+1}, s_2 s_{i+1}, \dots, s_{i+1}^2) = (s_1, \dots, s_{i+1}) s_{i+1} \\ &= (s_1, \dots, s_i) s_{i+1} = (s_1 s_{i+1}, \dots, s_i s_{i+1}) = z. \end{aligned}$$

This contradicts Definition 8.2. Hence $a_1 \subset a_2 \subset \dots$ is an infinite ascending chain contradicting the ascending chain condition.

THEOREM 9.1. *Let \mathfrak{S} be a semi-simple lattice. Then if each element of \mathfrak{S} can be expressed as a union of simple elements, \mathfrak{S} is a Boolean algebra.*

Proof. Let $a \in \mathfrak{S}$ have the representation

$$(9.1) \quad a = (s_1, \dots, s_n)$$

where s_1, \dots, s_k are distinct simple elements. The representation (9.1) is unique and s_1, \dots, s_k are the only simple elements which a divides. For let $a = (s_1, \dots, s_k) = (s'_1, \dots, s'_l)$. Multiplying by s'_i we have $s'_i = (s_1 s'_i, \dots, s_k s'_i)$. Hence all of the products are null except one, say $s_i s'_i$. Then $s_i s'_i = s'_i$ and hence $s_i \supset s'_i$ by Lemma 8.3. Thus $s_i = s'_i$ and $k = l$. If $a \supset s$, where s is simple and not equal to any of s_1, \dots, s_k , then $(s_1, s_2, \dots, s_k) = (s_1, \dots, s_k, s)$ contrary to the result we have just obtained.

We show now that the product of any two elements is equal to their crosscut.

We clearly have $[a, b] \supset ab$. Let $[a, b] = (s_1, \dots, s_k)$. Then since $a, b \supset [a, b]$, $a = (s_1, s_2, \dots, s_k, a')$ and $b = (s_1, \dots, s_k, b')$. Hence

$$ab = (s_1, \dots, s_k, a')(s_1, \dots, s_k, b') = (s_1, \dots, s_k, a'b') \supset [a, b].$$

Thus $[a, b] = ab$.

Since the product is distributive with respect to union, the crosscut must be distributive and hence \mathfrak{S} is distributive. Furthermore \mathfrak{S} is complemented. For let $a = (s_1, \dots, s_k)$, $u = (s_1, \dots, s_n)$ and define $a' = (s_{k+1}, \dots, s_n)$. Then $(a, a') = u$ and $[a, a'] = aa' = (s_1, \dots, s_n)(s_{k+1}, \dots, s_n) = z$. Hence \mathfrak{S} is a Boolean algebra.

In an arbitrary semisimple lattice, the set of elements which can be represented as a union of simple elements need not be closed with respect to crosscut as we shall show by an example. However, if we assume the modular* condition we have the following theorem.

THEOREM 9.2. *Let \mathfrak{S} be a modular semisimple lattice. Then the simple elements of \mathfrak{S} generate a Boolean algebra \mathfrak{S}_B . Moreover \mathfrak{S}_B is dense in \mathfrak{S} .*

Proof. Let U be the set of all elements of \mathfrak{S} which can be expressed as a

* For various statements of the modular axiom see Ore [4].

union of simple elements of \mathfrak{S} . The set U is obviously closed with respect to union. We shall show that U is dense in \mathfrak{S} and hence closed with respect to crosscut. Let $(s_1, \dots, s_n) \supset x$, and let $x \supset s_1, \dots, s_l, x \not\supset s_{l+1}, \dots, s_n$. Then $x = [x, (s_1, \dots, s_n)] = (s_1, \dots, s_l, [x, (s_{l+1}, \dots, s_n)])$ by the modular condition. If $[x, (s_{l+1}, \dots, s_n)] \neq z$, then there is a simple element s such that $[x, s_{l+1}, \dots, s_n] \supset s$. But then $x \supset s$ and $(s_{l+1}, \dots, s_n) \supset s$. Hence

$$s = s(s_{l+1}, \dots, s_n) = (s_{l+1}s, \dots, ss_n) = s_is$$

by Lemma 8.3. Thus $s = s_i$ and $x \supset s_i$ contrary to assumption. Hence $[x, (s_{l+1}, \dots, s_n)] = z$ and $x = (s_1, \dots, s_l)$.

Since U is dense in \mathfrak{S} , it is closed with respect to multiplication and is clearly semisimple. Moreover every element of U can be expressed as a union of simple elements. Hence by Theorem 9.1, $U = \mathfrak{S}_B$ is a Boolean algebra.

To show the significance of the modular condition in the previous theorem we give an example of a non-modular semisimple lattice in Fig. 4.

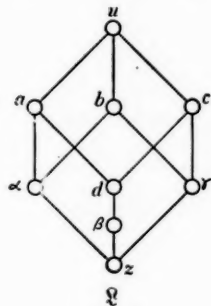


FIG. 4

If U denotes the set of elements of \mathfrak{Q} which can be expressed as a union of simple elements, we define a multiplication over \mathfrak{Q} as follows: If $x, y \in U$, $x \neq a$, $y \neq b$, then $xy = [x, y]$, $ac = \beta$, $dx = \beta$ or z according as $x \supset d$ or $x \not\supset d$. It can be readily verified that all of the multiplication postulates are satisfied. Also \mathfrak{Q} is non-modular since it contains the non-modular sublattice $\{a, \alpha, d, \beta, z\}$. The simple elements α, β, γ do not generate a Boolean algebra. In fact, U is not closed with respect to crosscut since $d = [(\alpha, \beta), (\beta, \gamma)]$.

THEOREM 9.3. *Let \mathfrak{S} be a modular semisimple lattice. Then if for each simple element s there exists an element $s' \neq u$ such that $(s, s') = u$, \mathfrak{S} is a Boolean algebra.*

Proof. We may take the s' 's to be divisor-free elements since if s'_i is not divisor-free, there exists a divisor-free element f_i such that $f_i \supset s'_i$. But then

$(s_i, f_i) \supset (s_i, s'_i) = u$. Let $v = (s_1, \dots, s_n)$. Then the length of chain from v to z is n . But now $[s'_1, s'_2, \dots, s'_n] = z$, since if $[s'_1, \dots, s'_n] \neq z$, there exists an s_i such that $[s'_1, \dots, s'_n] \supset s_i$. But then $s'_i \supset s_i$, which is impossible. Since $[s'_1, \dots, s'_n] = z$, the length of chain from u to z is equal to or less than n . But $u \supset v$. Hence $u = v$.

Theorem 9.3 gives immediately

THEOREM 9.4. *A complemented, modular, semisimple lattice is a Boolean algebra.*

We conclude with the statement of Theorem 9.3 in terms of the two-sided ideals of a non-commutative ring.

THEOREM 9.5. *Let R be a ring without radical in which the ascending and descending chain conditions hold for two-sided ideals. Then if for each two-sided ideal a there exists an ideal $a' \neq R$ such that $(a, a') = R$, R is a direct sum of two-sided simple ideals.*

Such an ideal a' always exists if a has a principle unit. For in that case we may take a' to be the set of all elements x such that $ax = 0$.

REFERENCES

1. G. Birkhoff, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 613-619.
2. R. P. Dilworth, Bulletin of the American Mathematical Society, vol. 44 (1938), pp. 262-267.
3. W. Krull, Mathematische Zeitschrift, vol. 28 (1928), pp. 481-503.
4. O. Ore, Annals of Mathematics, (2), vol. 36 (1935), pp. 406-432.
5. M. Ward, Annals of Mathematics, (2), vol. 39 (1938), pp. 558-568.
6. M. Ward and R. P. Dilworth, Proceedings of the National Academy of Sciences, vol. 24 (1938), pp. 162-164.
7. ———, these Transactions, vol. 45 (1939), pp. 335-354.

CALIFORNIA INSTITUTE OF TECHNOLOGY,
PASADENA, CALIF.

IDEAL THEORY AND ALGEBRAIC DIFFERENCE EQUATIONS*

BY

J. F. RITT AND H. W. RAUDENBUSH, JR.

Recent work of J. L. Doob, F. Herzog, W. C. Strodt and J. F. Ritt furnishes a theory of manifolds for systems of algebraic difference equations.† We present here a basis theorem for infinite systems of difference polynomials, and a restricted theory of ideals; there is obtained thus a counterpart, for difference equations, of Raudenbush's work on differential equations.‡

In the theory of algebraic polynomials, one derives an infinite system from a basis by forming linear combinations. A system of differential polynomials is obtained from a basis by differentiations, linear combinations and the extraction of roots. For difference polynomials, one performs "shufflings" in succession, each shuffling consisting in taking *transforms*, performing linear combinations and factoring forms into products of transforms.

We leave open the question as to how many shufflings are necessary in order to produce a system from its basis. Conceivably, an infinite number may be necessary in certain cases; we give an example, in §15, for which two shufflings are required.

We shall, in this paper, work with difference polynomials whose coefficients lie in an abstract field. If we have decided to relinquish the meromorphic coefficients used in R. D., it is because, with the present undeveloped state of the analytic theory of nonlinear difference equations, it appears tactical to proceed in an algebraic direction, hoping for analytical developments to follow. For example, the establishment of existence theorems which will permit the translation into analytic terms of the *Nullstellensatz* presented in §13, is a problem with a distinct challenge.

DIFFERENCE RINGS

1. Let \mathcal{R} denote a commutative ring§ possessing a unit element. Let us

* Presented to the Society, September 8, 1939; received by the editors May 24, 1939.

† For references, see *Semicentennial Addresses of the American Mathematical Society*, New York, 1938, pp. 54, 55. The present paper attaches particularly to Ritt and Doob, *Systems of algebraic difference equations*, American Journal of Mathematics, vol 55 (1933). That paper will be designated below by R. D.

‡ Actually, our considerations hold for equations involving a substitution of any type; for instance, for q -difference equations.

§ Defined as in van der Waerden, *Moderne Algebra*, chap. 3. Small italics will, until §9, usually represent elements of \mathcal{R} .

suppose that, for every element a of \mathcal{R} , \mathcal{R} contains a unique element a_1 , called the *transform* of a , the correspondence between elements and their transforms being such that:

- (α) the transform of unity is unity;
- (β) for every a and b in \mathcal{R} , $(a+b)_1 = a_1 + b_1$ and $(ab)_1 = a_1 b_1$.

We shall, under these circumstances, call \mathcal{R} a *difference ring*. If \mathcal{R} , in addition to being a ring, is a field, we shall call \mathcal{R} a *difference field*.*

In everything which follows, we deal with a fixed difference ring \mathcal{R} .

2. We denote $(a_1)_1$ by a_2 and, by induction, define a_n as $(a_{n-1})_1$ for every $n > 1$. We shall call a_n the *transform of a of order n* or the *n th transform of a* . The element a will be described as its own transform of order 0 and will be denoted, at times, by a_0 . We shall refer to the a_n , $n=0, 1, \dots$, as *transforms of a* ; the transform of a will continue to mean, as above, a_1 .

IDEALS

3. An ideal π contained in \mathcal{R} will be called a *difference ideal* if, given any element a in \mathcal{R} , the presence in π of either of a and a_1 , where a_1 is the transform of a , implies the presence in π of the other. Thus, if a is in π , π contains every transform of a and also contains every element in \mathcal{R} of which a is a transform of some order.†

4. A difference ideal π will be called *perfect* if, whenever a is such that some product of positive integral powers of transforms of a is contained in π , a is also contained in π . That is, if

$$a_p a_q^i \cdots a_r^k,$$

where p, q, \dots, r are distinct nonnegative integers and i, j, \dots, k are positive integers, is in π , a is in π .

A difference ideal π will be called *prime* if, whenever ab is in π , at least one of a and b is in π ; π will thus be a prime ideal in the sense in which that term is regularly used in algebra. *Every prime difference ideal is perfect.*

Henceforth, unless other indications are given, *ideal* will mean difference ideal.

* Let $b \neq 0$. By (α), $b_1(1/b)_1 = 1$. Hence $b_1 \neq 0$ and $(a/b)_1 = a_1/b_1$ for every a .

† Our insistence that π contain, together with a , all transforms of a "of negative orders" which may exist in \mathcal{R} , is explained by the material of §4. Our definition of ideal appears to have sufficient generality for the purposes of the concrete applications; one will notice, for instance, that a difference form with meromorphic coefficients has the same manifold as its transform. (Cf. R. D.) The fact that many investigations on difference equations deal with half of the complex plane seems to make it undesirable to assume that, for every a in \mathcal{R} , there is an element in \mathcal{R} of which a is a transform.

PERFECT IDEALS GENERATED BY A SET OF ELEMENTS

5. Let σ be any set of elements in \mathcal{R} .^{*} There exist perfect ideals in \mathcal{R} , for instance \mathcal{R} itself, which contain σ . The intersection of all such perfect ideals is a perfect ideal which contains σ . We denote this intersection by $\{\sigma\}$ and call it the *perfect ideal generated by σ* .

We shall study the relationship of $\{\sigma\}$ to σ .

Let τ be any set of elements of \mathcal{R} . Consider all elements of \mathcal{R} which are of the form

$$au + bv + \cdots + cw,$$

where u, v, \dots, w are transforms of any orders of elements of τ and a, b, \dots, c are in \mathcal{R} . The totality of such elements will be denoted by $[\tau]$; $[\tau]$ may not be a difference ideal, but it will be an ideal in the sense of algebra and it will be closed with respect to "transforming."

Again, let a be any element in \mathcal{R} which is such that some product of positive powers of transforms of a is in τ . The totality of such elements a will be denoted by τ' .[†]

Returning to σ above, let $\sigma_1 = [\sigma]'$ and, continuing inductively, let $\sigma_n = [\sigma_{n-1}]'$ for every $n > 1$. The logical sum, or what is the same, the *limit*, of the sets σ_n is easily seen to be $\{\sigma\}$.

Because $[\sigma]$ and the $[\sigma_n]$ are closed with respect to "transforming," every σ_n is also so closed. Thus, for $n \geq 1$, each element of $[\sigma_n]$ is a linear combination of elements of σ_n .

In what follows, a plus sign between two sets will indicate that the logical sum of the sets is to be taken.

6. We prove the following:

LEMMA I. *Let σ be any set of elements of \mathcal{R} and a and b any two elements of \mathcal{R} . If d is contained in $(\sigma + a)_n$ and e in $(\sigma + b)_n$, $n \geq 1$, then de is contained in $(\sigma + ab)_{n+1}$.*[‡]

First, let $n = 1$. There exist a product \bar{d} of positive powers of transforms of d , and an \bar{e} similarly related to e , which have expressions

$$\begin{aligned}\bar{d} &= gu + \cdots + hv + ka_i + \cdots + la_j, \\ \bar{e} &= g'u' + \cdots + h'v' + k'b_p + \cdots + l'b_q,\end{aligned}$$

^{*} For the purposes of §§11, 12, it is desirable to allow a given element of \mathcal{R} to occur more than once in σ . Thus, the elements in σ are supposed to be provided with marks and a single element of \mathcal{R} may appear many times in σ , each time in association with a different mark. When we have to do with ideals, however, a given element will be assumed to appear only once.

[†] In $[\tau]$ and in τ' a given element will be understood to occur only once. The notation in the present paragraphs, as regards accents and subscripts, is of an episodic character.

[‡] The parentheses are ordinary symbols of aggregation. Thus, $(\sigma + a)_1 = [\sigma + a]'$.

where u, \dots, v and u', \dots, v' are transforms of elements of σ and the subscripted a and b are transforms of a and b . Thus $\bar{d}\bar{e}$ has an expression in which some terms are in $[\sigma]$ and in which the others are of the type $f a_r b_s$. For any r and s , the product of $a_r b_s$ by a suitable transform of itself is a multiple* of a transform of ab . Thus every $a_r b_s$ is in $[ab]'$ and, a fortiori, in $(\sigma+ab)_1$. Then $\bar{d}\bar{e}$ is in $[(\sigma+ab)_1]$. Some product of powers of transforms of de is a multiple of $\bar{d}\bar{e}$. Thus de is in $(\sigma+ab)_2$.

Now, let $n=2$. Let \bar{d} , described as above, be in $[(\sigma+a)_1]$. By §5, \bar{d} is a linear combination of elements of $(\sigma+a)_1$. We use an \bar{e} , described as above, which is linear in elements of $(\sigma+b)_1$. Then $\bar{d}\bar{e}$ has an expression in which each term is of the type guv with u in $(\sigma+a)_1$ and v in $(\sigma+b)_1$. Now uv , by the case of $n=1$, is in $(\sigma+ab)_2$. Hence $\bar{d}\bar{e}$ is in $[(\sigma+ab)_2]$. This puts de in $(\sigma+ab)_3$.

The proof continues by induction.

LEMMA II. *Let σ be any set of elements of \mathcal{R} and a and b any two elements of \mathcal{R} . Then $\{\sigma+ab\}$ is the intersection of $\{\sigma+a\}$ and $\{\sigma+b\}$.*

We have only to show that, c being any element in the intersection, c is contained in $\{\sigma+ab\}$. Let n be such that c is contained in $(\sigma+a)_n$ and in $(\sigma+b)_n$. Then c^2 is in $(\sigma+ab)_{n+1}$. Thus c is also in $(\sigma+ab)_{n+1}$.

BASES

7. Let σ be a system of elements in \mathcal{R} . A finite subset ϕ of σ will be called a *basis* of σ if $\{\phi\}$ contains σ .

A finite system of elements is a basis for itself. If every infinite system of elements in \mathcal{R} has a basis, we shall call \mathcal{R} a *difference ring with a basis theorem*.

DECOMPOSITION OF PERFECT DIFFERENCE IDEALS

8. Let \mathcal{R} have a basis theorem. We prove the theorem:

THEOREM. *Every perfect ideal in \mathcal{R} is the intersection of a finite set of prime ideals.*

Let π be a perfect ideal for which our statement is false. Then π is not prime. Let ab be in π while neither a nor b is. Then π is the intersection of $\{\pi+a\}$ and $\{\pi+b\}$ (Lemma II). At least one of the two latter ideals does not have the property of being the intersection of a finite set of prime ideals. Of the two ideals, let π_1 designate one which lacks the property. We give π_1 the treatment accorded to π and continue, forming a sequence of perfect ideals

$$(1) \quad \pi, \pi_1, \dots, \pi_n, \dots,$$

* The meaning is obvious.

each a proper part of its successor. Let σ be the logical sum of the ideals in (1) and let ϕ be a basis of σ . There is some π_n which contains ϕ . That π_n will contain σ . This contradiction proves the theorem.

It is easy now to see that every perfect ideal in \mathcal{R} has a *unique* representation as the intersection of a finite number of prime ideals *none of which contains any other*.

IDEALS OF DIFFERENCE POLYNOMIALS

9. Let n be any positive integer. We consider n symbols $y_1(x), \dots, y_n(x)$. If j is any nonnegative integer, we shall call the symbol $y_i(x+j)$ the *jth transform* of $y_i(x)$.

Let \mathcal{F} be a given difference field. By a *difference polynomial* we shall mean a polynomial in a certain (eo ipso finite) number of the $y_i(x+j)$, with coefficients in \mathcal{F} . As a rule, we shall substitute the briefer term *form* for "difference polynomial." By the *transform* of a form A , we mean the form obtained when x is replaced by $(x+1)$ in the $y_i(x+j)$ appearing in A , and when the coefficients in A are replaced by their transforms. Transforms of higher order are defined similarly. Because the transform of unity is unity, these definitions are consistent with the definition given above of the *jth transform* of $y_i(x)$.

The totality of forms with coefficients in \mathcal{F} is a difference ring, which we shall call the *ring of forms in the unknowns* y_1, \dots, y_n . Any form of this ring will be called a *form in* y_1, \dots, y_n .

10. We prove the theorem:

THEOREM. *For any difference field \mathcal{F} , the ring of difference polynomials in the unknowns y_1, \dots, y_n is a difference ring with a basis theorem.*

We assume the theorem to be false and work towards a contradiction. We shall use methods and results of R. D. The items of that paper which will be employed here acquire validity, with no essential change, for an abstract field \mathcal{F} .

11. We prove the following lemma:

LEMMA III. *Let Σ be a system of forms in y_1, \dots, y_n which has no basis. Let F_1, \dots, F_s be such that, by multiplying each form in Σ by some product of nonnegative powers of transforms of F_1, \dots, F_s , a system Λ is obtained which has a basis. Then at least one of the systems $\Sigma + F_i$, $i=1, \dots, s$, has no basis.*

Suppose that every $\Sigma + F_i$ has a basis. Then, for each i , there is a finite subset Φ_i of $\Sigma + F_i$ such that $\{\Phi_i\}$ contains $\Sigma + F_i$. As Φ_i may evidently be replaced by any finite subset of $\Sigma + F_i$ which contains Φ_i , we may (and shall) suppose Φ_i , for every i , to be of the type

$$(2) \quad F_i, A_1, \dots, A_q,$$

with the set

$$(3) \quad A_1, \dots, A_q$$

independent of i . Enlarging (3) if necessary, we assume that the subset of Λ obtained from (3) by the above described multiplications is a basis of Λ . Thus, the perfect ideal generated by the set (3) contains Λ .

Let L be any form in Σ . Then L is contained, for every i , in the perfect ideal generated by the set (2). Certainly, L is contained in the perfect ideal generated by

$$L, A_1, \dots, A_q.$$

By Lemma II of §6, L is contained in the perfect ideal generated by

$$(4) \quad F_1 F_2 \cdots F_s L, A_1, \dots, A_q.$$

Some KL , with K a product of powers of transforms of the F_i , belongs to Λ . Now an appropriate product of powers of transforms of $F_1 F_2 \cdots F_s L$ is a multiple of KL . Since Λ is contained in the perfect ideal generated by (3), $F_1 \cdots F_s L$ is also so contained. Inspecting (4), we see that L is contained in the perfect ideal generated by (3). Then (3) is a basis of Σ . This contradiction proves the lemma.

12. From among all systems of forms in y_1, \dots, y_n which have no basis, we select one, Σ , whose basic sets are not higher than those of any other system which has no basis. Let

$$(5) \quad A_1, \dots, A_r$$

be a basic set of Σ . Then A_1 must be of class greater than 0, else Σ would be contained in $\{A_1\}$. Let I_i be the initial of A_i , $i=1, \dots, r$.

For every form of Σ not in (5), let a remainder with respect to (5) be found. Let Λ be the system composed of the forms in (5) and of the products of the forms of Σ not in (5) by the power products of transforms of the I_i used in forming the remainders. Let Ω be the system composed of (5) and of the remainders of the forms of Σ not in (5).

By the considerations of R. D., Ω has a basis. Such a basis is a basis of $\{\Omega\}$. Now $[\Lambda]$ is identical with $[\Omega]$, so that $\{\Lambda\}$ is identical with $\{\Omega\}$. Thus $\{\Lambda\}$ has a basis. It is easy now to see that $\{\Lambda\}$ has a basis composed of forms of Λ . Such a basis is a basis of Λ .

The lemma of §11 informs us now that some $\Sigma + I_i$ has no basis. As in R. D., this is impossible. The theorem is proved.

HILBERT-NETTO THEOREM

13. Following the procedure of Raudenbush for differential equations, one can prove that if Φ is a finite system of difference polynomials with coefficients in a difference field \mathcal{F} and if G is a form which is not in $\{\Phi\}$, then there exists an extension \mathcal{F}' of \mathcal{F} , containing a solution of Φ which is not a solution of G .*

DERIVATION OF AN INFINITE SYSTEM FROM A BASIS

14. Let Σ be an infinite system of forms in y_1, \dots, y_n and Φ a basis of Σ . Given any form A in Σ , there is some Φ_i (notation as in §5) which contains A . It is natural to ask whether there is some Φ_i which contains Σ , and, indeed, whether Φ_1 contains Σ .

We shall present a system Σ which has no basis Φ for which Φ_1 holds Σ . Whether there is a system with no basis for which Φ_2 contains Σ , we do not know.

15. Our example will deal with forms in a single unknown y . We use the field of constants, each constant being its own transform. The j th transform of y will be denoted by y_j . The system Σ will consist of the sequence of forms A_0, A_1, \dots , where

$$A_j = (y_0 + y_1 + \dots + y_{2^j-1})(y_{2^j} + \dots + y_{2^{j+1}-1}).$$

Let m be any nonnegative integer. Let Φ consist of A_0, \dots, A_m . We shall show that A_{m+1} is not contained in Φ_1 .

Consider any A_j with $j \leq m$. Each term in A_j is of the type $y_a y_b$ with

$$(6) \quad 0 < b - a \leq 2^{m+1} - 1.$$

Hence, if G is any form in $[\Phi]$, each term in G has, among its letters, two letters y_a and y_b where a and b satisfy (6). We shall show that if K is a product of powers of transforms of A_{m+1} , K contains a term each of whose letters y_j has its subscript j divisible by 2^{m+1} . This will prove that A_{m+1} is not contained in Φ_1 .

Let B_i represent the i th transform of A_{m+1} , $i=0, 1, \dots$. Let B_j^r , with $r>0$, be one of the powers of which K is a product. Let

$$j = p2^{m+1} + q,$$

with p and q nonnegative integers and $q < 2^{m+1}$. We have

$$(7) \quad B_j = (y_j + \dots + y_{2^{m+1}+j-1})(y_{2^{m+1}+j} + \dots + y_{2^{m+2}+j-1}).$$

* These Transactions, vol. 36 (1934), p. 368. The term *extension* is self-explanatory.

If $q=0$, the first parenthesis in (7) contains $y_{p2^{m+1}}$ and the second contains $y_{(p+1)2^{m+1}}$. If $q>0$, the first parenthesis contains $y_{(p+1)2^{m+1}}$ and the second $y_{(p+2)2^{m+1}}$. In any case, B_j contains one and only one term $y_a y_b$ in which a and b are both divisible by 2^{m+1} . Furthermore, B_j contains a term which is a power of $y_a y_b$, and that term of B_j is the only one in which every subscript is divisible by 2^{m+1} . Our statement with respect to K follows.

We observe that A_0 is a basis of Σ and that, if Ψ represents A_0 , Ψ_1 contains y_0 so that $[\Psi_1]$ and Ψ_2 contain Σ .

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.,
QUEENS COLLEGE,
FLUSHING, N. Y.

EXPONENT TRAJECTORIES IN SYMBOLIC DYNAMICS*

BY

RUFUS OLDENBURGER

1. **Introduction.** Morse, Hedlund,† and others have developed the theory of dynamics from the symbolic point of view. This theory is concerned in the main with the periodicity, recurrence, and transitivity properties of symbolic trajectories and rays. Morse has made use of exponents on symbols. Unless a trajectory T is of a very special type, it can be shown that the exponents on the symbols in a symbolic trajectory T form a symbolic trajectory T_e termed the "exponent trajectory" of T . The trajectory T_e is uniquely determined by T . Similar considerations hold for rays. In the present paper we are concerned with relations between a trajectory or ray and the associated exponent trajectory or ray. In particular we prove that a periodic or recurrent trajectory T has a periodic or recurrent exponent trajectory T_e respectively, while a transitive ray R has an exponent ray R_e which is in a sense also transitive. Further, if a trajectory T is periodic, T is distinct from its exponent trajectory. There exist, however, trajectories identical with their exponent trajectories, and in the case of trajectories generated by the symbols 1, 2 only, there is *one and only one such trajectory*. The term "identical" is used here in the usual sense, and will be defined explicitly in the next section. In the paper referred to above, Morse and Hedlund have given some methods of constructing recurrent trajectories from a given recurrent trajectory. The introduction of exponent trajectories yields another method of constructing such trajectories. Whether or not there exist recurrent trajectories identical with their exponent trajectories is still an open question.

2. **Definitions and conventions.** We shall use the term "symbolic trajectory" in a slightly more general sense than that employed by Morse and Hedlund in that we shall allow an infinite set of generating symbols. Let S_1 denote a sequence $abc \dots$ of symbols a, b, c, \dots which may or may not be taken from a finite set of distinct symbols, and let S_2 denote a second such sequence $\alpha\beta\gamma \dots$. Let S_2^{-1} denote the sequence $\dots \gamma\beta\alpha$ of symbols obtained from S_2 by reversing the order of the symbols in S_2 . The sequence $S_2^{-1}S_1$, given by $\dots \gamma\beta\alpha abc \dots$, is termed a *symbolic trajectory*, or simply a *trajectory*. The sequence S_1 (also S_2^{-1}) is termed a *ray*. The symbol a in S_1 is termed the *initial symbol* of the ray S_1 . We shall have occasion to use the notation

* Presented to the Society, October 28, 1939; received by the editors July 24, 1939.

† Marston Morse and Gustav A. Hedlund, *Symbolic dynamics*, American Journal of Mathematics, vol. 60 (1938), pp. 815-866.

$S_1 = abc \dots$ meaning that S_1 is the sequence $abc \dots$. A finite sequence $ab \dots k$ of symbols is termed a *block*. If there are n symbols in the set a, b, \dots, k , the block $ab \dots k$ is said to be of *length* n , and will be called an *n-block*. If B is a block, the length of B will be denoted by $l(B)$. We shall write $B = ab \dots k$ to indicate that B is the block $ab \dots k$. If $B_1 = a_1 \dots a_m$, $B_2 = b_1 \dots b_n$, then $B_1 B_2$ is the block $a_1 \dots a_m b_1 \dots b_n$. The blocks B_1 and B_2 are the same if $m = n$ and the symbol a_i is identical with the symbol b_i for each i in the range $1, 2, \dots, n$. In a block $C = a_{-n} \dots a_{-1} a_0 a_1 \dots a_n$ of odd length, we term a_0 the *central symbol* of C . A trajectory T can be written as

$$\dots a_{-2} a_{-1} a_0 a_1 a_2 \dots$$

The symbols a_i and a_j are said to be in *different positions* in T if $i \neq j$. If $i = j$ these elements are in the *same position* in T . Let a_0 denote a symbol in a fixed position in a trajectory T_1 . The trajectory T_1 is said to be *identical* with a trajectory T_2 if T_2 contains the symbol a_0 in a fixed position so that for each n the block A_n in T_1 of length $2n+1$ containing a_0 as central symbol is identical with the $(2n+1)$ -block B_n of T_2 containing a_0 as central symbol.

Sequences of consecutive symbols of a trajectory T (or ray R or block B) which form a block or ray we term a *subblock* or *subray* of T (or R or B), and they are said to be *contained* in T (or R or B). As remarked above the symbols in a trajectory T (or ray R or block B) are taken from a finite or infinite set S of distinct symbols, which will be termed the *generating symbols* of T (or R or B). A block $a \dots a$ formed by repeating the symbol a n times is written as a^n . The symbol n in a^n is termed the *exponent* of a in a^n , and a is termed the *base* in a^n . We term a^n a *power*. We write a block B as a sequence of powers such that the bases in consecutive powers are distinct. The exponents then form the *exponent block* B_e of B . Unless a trajectory T contains a subray formed by only one generating symbol, T can be written as a sequence

$$(1) \quad \dots a^p b^q c^r \dots,$$

where no two consecutive bases are identical. The exponents in (1) form a trajectory $\dots pqr \dots$, which we term the *exponent trajectory* T_e of T . Similarly, if a ray R does not contain a subray formed by one generating symbol, the ray R can be written as $a^p b^q c^r \dots$, where consecutive bases are distinct. The exponents then form the *exponent ray* R_e of R . A trajectory T (or ray R) will be termed *admissible* if it has an exponent trajectory (or ray); that is, T (or R) does not contain a subray of the form $aaa \dots$ or $\dots aaa$.

A trajectory T is *periodic* if it can be written as a sequence

$$(2) \quad \dots BBB \dots$$

of blocks identical with a block B . If B is a block of shortest length such that T can be written as (2), the block B is said to be a *period block* of T , and its length is termed the *period* of T . A trajectory T is termed *recurrent* if for each n there exists an m such that each block of length n in T is contained in each m -block of T . If T is recurrent, for each n there exists a least m such that each m -block of T contains each n -block of T . We write $R(n) = m$, and term $R(n)$ the *recurrency function* of T . A ray R is said to be *transitive* if every possible block that can be formed from the generating symbols of R is a sub-block of R .

3. **Periodicity, recurrence, and transitivity of exponent trajectories.** We shall now prove the following theorem.

THEOREM 1. *If a trajectory T in two or more generating symbols is periodic, T is admissible and the exponent trajectory T_e is periodic.*

Let B represent a period block of T so that T is given by (2). Suppose that B begins with the symbol a and is preceded by a in T . Then B is of the form $a^r b^s \cdots c^t a^u$, where no two consecutive symbols in the set a, b, \cdots, c, a are identical. The block $C = a^w b^s \cdots c^t$, where $w = u + r$, is then also a period block of T . The block $C_e = ws \cdots t$ thus occurs in T_e , and T_e is of the form $\cdots C_e C_e C_e \cdots$, whence the theorem is proved.

THEOREM 2. *The exponent trajectory T_e of an admissible periodic trajectory T is distinct from T .*

As noted above T contains a period block $C = a^w b^s \cdots c^t$, where $a \neq c$, and the exponent block $C_e = ws \cdots t$ of C is a subblock of T_e . Evidently C_e or a subblock of C_e is a period block of T_e . The period of T is $\omega = w + s \cdots + t$. The period of T_e is no greater than the length L of C_e . If at least one of the symbols in C_e is greater than 1, we have $\omega > L$. If all of the symbols in C_e equal 1, the period of T_e is 1 and certainly less than ω .

Morse and Hedlund* have exhibited a nonperiodic recurrent trajectory T in four symbols with the property that consecutive symbols in T are distinct. It follows that in this case T_e is of the form

$$(3) \quad \cdots 111 \cdots$$

Since (3) is periodic, there exist nonperiodic trajectories whose exponent trajectories are periodic. That this is not true of trajectories with two generating symbols is stated in the theorem which follows.

THEOREM 3. *An admissible trajectory T with two generating symbols is periodic if and only if its exponent trajectory T_e is periodic.*

* See the reference to Morse and Hedlund above, p. 844.

Let the generating symbols be denoted by 1, 2. Let the period of T_e be denoted by ξ , and a period block of T_e by $B_e = a_1 \cdots a_\xi$. Let B be a block of T with exponent block B_e . If ξ is even, the first and last symbols of B are distinct, for $B = 1^{a_1} 2^{a_2} 1^{a_3} \cdots 2^{a_\xi}$ or $2^{a_1} 1^{a_2} 2^{a_3} \cdots 1^{a_\xi}$. Hence T is given by (2), and T is periodic. If ξ is odd, the first and last symbols of B are identical. It follows that T is given by

$$(4) \quad \cdots B_1 B_2 B_1 B_2 B_1 \cdots,$$

where $B_1 = 1^{a_1} 2^{a_2} 1^{a_3} \cdots 1^{a_\xi}$, $B_2 = 2^{a_1} 1^{a_2} 2^{a_3} \cdots 2^{a_\xi}$. Hence T is periodic.

THEOREM 4. *If the exponent trajectory T_e of a periodic trajectory T in two generating symbols has the period block $a_1 a_2 \cdots a_\xi$, the trajectory T has the period ω , where*

$$(5) \quad \omega = \sum_{j=1}^{\xi} a_j,$$

$$(6) \quad \omega = 2 \left(\sum_{j=1}^{\xi} a_j \right),$$

according as ξ is even or odd.

From the proof of Theorem 3 it follows in the case where ξ is even that the trajectory T is given by (2), where $B = 1^{a_1} 2^{a_2} 1^{a_3} \cdots 2^{a_\xi}$ or $2^{a_1} 1^{a_2} 2^{a_3} \cdots 1^{a_\xi}$. Hence $\omega \leq (a_1 + \cdots + a_\xi)$. It is no restriction to suppose that $B = 1^{a_1} 2^{a_2} 1^{a_3} \cdots 2^{a_\xi}$. If B is not a period block, a subblock $1^{a_1} 2^{a_2} 1^{a_3} \cdots 2^{a_j}$, $j < \xi$, of B is a period block of T . Then T_e is given by $\cdots B'_e B'_e B'_e \cdots$, where $B'_e = a_1 a_2 \cdots a_j$. The trajectory T_e thus has a period less than ξ , which is impossible. It follows that (5) is valid.

It follows from the proof of Theorem 3 that if ξ is odd the period of T is not greater than $2(a_1 + a_2 + \cdots + a_\xi)$, and that T is given by (4). By the argument of the preceding paragraph the period of T cannot be less than $(a_1 + a_2 + \cdots + a_\xi)$. Hence T has the period block $1^{a_1} 2^{a_2} \cdots 1^{a_\xi} 2^{a_1} 1^{a_2} \cdots 2^{a_j}$ or the equivalent block with the symbols 1 and 2 interchanged. Thus T_e is given by $\cdots B''_e B''_e B''_e \cdots$, where $B''_e = a_1 a_2 \cdots a_\xi a_1 a_2 \cdots a_j$. Since ξ divides the length of B_e , we have $j = \xi$, whence (6) is valid.

From Theorem 4 it is evident that the number of periodic trajectories of period ω with two generating symbols is the number of solutions of

$$\sum_{j=1}^{2n} a_j = \omega, \quad 2 \left(\sum_{j=1}^{2n+1} a_j \right) = \omega,$$

where the a 's and n are integers, and the blocks $a_1 a_2 \cdots a_\xi$ ($\xi = 2n, 2n+1$) are not of the form $DD \cdots D$, that is, formed by the repetition of a block.

LEMMA 1. *A recurrent trajectory T with two or more generating symbols is admissible.*

Since T contains a block ab , where a and b are distinct, and this block cannot be contained in a subray with one generating symbol, it follows that each exponent is finite, and the exponent trajectory T_e exists.

LEMMA 2. *If an admissible trajectory T is recurrent, its exponent trajectory contains a finite number of generating symbols.*

Consider again a subblock ab of T where $a \neq b$. If there exists in T a sequence of blocks $a_1^{n_1}, a_2^{n_2}, \dots$, where the sequence n_1, n_2, \dots is unbounded, then there exists an arbitrarily long block which does not contain ab . Hence T is not recurrent. Thus Lemma 2 is proved.

THEOREM 5. *If an admissible trajectory T is recurrent, the exponent trajectory T_e of T is recurrent.*

Consider a block $B_e = ps \dots q$ of T_e . There is a corresponding block $B = a^p b^s \dots c^q$ of T bordered on the left and right by symbols g and h respectively, where $g \neq a$ and $h \neq c$. Since T is recurrent, the block gBh occurs in each block of T of length $R(n)$, where n is the length of gBh , and $R(n)$ is the recurrency function of T . Thus in each subblock B' of T of length $R(n)$ there occurs a block $g^\alpha B h^\beta$, where $\alpha \geq 1$, and $\beta \geq 1$. Each block B' is contained in a block B'' , where B'' is preceded in T by a symbol distinct from the first symbol of B'' , and followed by a symbol distinct from the last symbol of B'' , and the exponent block of B'' has the same length as the exponent block of B' . Evidently, B_e is contained in the block of exponents of each block B'' . Let t be the maximum length of the exponent blocks of the blocks of type B'' . We denote the exponent block of a block B'' by B_e'' . Each exponent block C_e in T_e of length t corresponds to a block C of T which contains a block B'' as subblock. It follows that each block of T_e of length t contains B_e . Let r denote the length of B_e . There are a finite number of blocks $B_{e1}, B_{e2}, \dots, B_{ep}$ in T_e of length r . There exist numbers t_1, t_2, \dots, t_p such that for each i ($i = 1, 2, \dots, p$) B_{ei} is contained in each t_i -block of T_e . Let $R_e(r)$ denote the maximum of the numbers t_1, t_2, \dots, t_p . Then each r -block of T_e is contained in each $R_e(r)$ -block. Thus T_e is recurrent.

COROLLARY 1. *If T is a recurrent nonperiodic trajectory in two generating symbols, the exponent trajectory T_e of T is a recurrent nonperiodic trajectory.*

It is obvious that a non-recurrent trajectory T may have a recurrent exponent trajectory T_e . It is necessary even in the case of two generating symbols to impose an additional restriction on T_e to insure the recurrence of T . We shall give the additional restriction for the case of two generating symbols.

We say that a trajectory T in two generating symbols is *strongly recurrent* if for each n and n -block B in T there exists an integer $R(n)$ such that if B_1, B_2 are any nonoverlapping blocks of length $R(n)$, the block B_1 contains a block B whose first symbol is separated from the first symbol of a block B in B_2 by an odd number of symbols. An immediate result is the following theorem.

THEOREM 6. *An admissible trajectory T in two generating symbols is recurrent if and only if its exponent trajectory is strongly recurrent.*

Certain inequality relations exist between the recurrency function of a recurrent trajectory T and the recurrency function of the exponent trajectory T_e of T . For the sake of brevity these relations will be omitted.

That the following theorem is true appears from the definition of transitivity.

THEOREM 7. *A transitive ray in two or more generating symbols is admissible.*

THEOREM 8. *The exponent ray R_e of a transitive ray R in two or more generating symbols is transitive.*

It is evident that R_e has the infinite set $1, 2, 3, \dots$ of generating symbols. We denote this set by S . Let l, m, n, \dots, p be an arbitrary subset of S containing μ symbols not necessarily distinct, and let q, r, s, \dots, t be a second subset of μ symbols in S not necessarily distinct. By assumption R contains at least two distinct generating symbols a, b . Since R is transitive, R contains the block $\alpha B \gamma$, where

$$B = a_1^{l_1} a_2^{l_2} \dots a_q^{l_q} a_{q+1}^{m_1} \dots a_{q+r}^{m_r} a_{q+r+1}^{n_1} \dots a_{q+r+s}^{n_s} \dots a_{q+r+s+\dots+u+1}^{p_1} \dots a_{q+r+s+\dots+u+t}^{p_t}$$

$\alpha \neq a_1, \gamma \neq a_{q+r+s+\dots+u+t}$, the exponent block of B is $B_e = l^q m^r n^s \dots p^t$, and the a 's are alternately equal to a and b so that $a_1 = a, a_2 = b, a_3 = a, \dots$. Thus R_e contains each block B_e that can be formed from the symbols in S , whence R_e is transitive.

Theorem 8 can be extended to "transitive trajectories" with no subray generated by one symbol only.

4. A trajectory identical with its exponent trajectory. In Theorem 2 we noted that a periodic trajectory is distinct from its exponent trajectory. That this is not true for trajectories in general is a consequence of the theorem which follows.

THEOREM 9. *There exists a trajectory identical with its exponent trajectory.*

We let B_0 denote the block 212, and let $B_1 = 2$. We form the trajectory

$$(7) \quad \dots B_3^{-1} B_2^{-1} B_1^{-1} B_0 B_1 B_2 B_3 \dots,$$

where B_i is the exponent block of B_{i+1} for each $i > 0$, the last symbol of B_i is distinct from the first symbol of B_{i+1} for each $i > 0$, and B_i^{-1} denotes the block obtained from B_i by reversing the symbols in B_i . We illustrate by giving some of the blocks B_i explicitly:

$$B_2 = 11, B_3 = 21, B_4 = 221, B_5 = 22112, B_6 = 11221211.$$

We note that for $i > 0$ the block B_i^{-1} is the exponent block of B_{i+1}^{-1} . Thus (7) is the sequence

$$(8) \quad \dots, 21122, 122, 12, 11, 2, 212, 2, 11, 21, 221, 22112, \dots,$$

where we have separated the blocks B_i and B_i^{-1} by commas. The exponent block of $B_1^{-1} B_0 B_1$ is B_0 . From this statement and the definition of (7), it appears that the exponent block of $B_r^{-1} \dots B_2^{-1} B_1^{-1} B_0 B_1 B_2 \dots B_r$ is the block $B_{r-1}^{-1} \dots B_1^{-1} B_0 B_1 \dots B_{r-1}$. Thus (7) has an exponent trajectory and is identical with it.

Employing the same technique as that used in constructing (7) and using more than two symbols, one can construct an unlimited number of trajectories identical with their exponent trajectories. We shall prove later the uniqueness of (7) for the class of trajectories in two generating symbols 1, 2.

5. Proper exponent blocks and join-blocks in trajectories with generating symbols 1, 2. Consider an arbitrary subblock B of a trajectory T in generating symbols 1, 2 where the exponent trajectory T_e of T contains the same generating symbols. The block B has an exponent block B_e which does not necessarily occur as a subblock of the exponent trajectory T_e of T since B may be preceded by or followed by a symbol identical with the first or last symbol of B respectively. For this reason we associate with B a new type of exponent block. Consider the block B_e of exponents of B which occur in T_e and can be determined without reference to T from B alone and the fact that the exponents equal 1 or 2. We term B_e the *proper exponent block* of B . We similarly speak of a *proper exponent ray*. We let C_1, C_2 be consecutive subblocks of the trajectory T so that $C_1 C_2$ is a subblock of T . We denote the proper exponent blocks of C_1 and C_2 by D_1 and D_2 respectively. The proper exponent block of $C_1 C_2$ is a block $D_1 J D_2$. We shall say that J is the *exponent block due to the join* of C_1 and C_2 . Obviously, J is either vacuous, or is one of the blocks 1, 2, or 11.

THEOREM 10. Let T_e be the exponent trajectory of a trajectory T , and suppose that T and T_e have the same generating symbols 1, 2. The length of the proper exponent block B_e of a block B in T satisfies the formula

$$(9) \quad L(B_e) \leq L(B) - 2$$

if $B \neq \alpha$, α^2 ($\alpha = 1, 2$). If B has an intermediate block 1^2 or 2^2 , then

$$(10) \quad L(B_e) \leq L(B) - 3.$$

In any case $L(B_e) \leq L(B) - 1$. We write $B = a_1 a_2' a_3'' \cdots a_{n-1}' a_n$, $n \geq 2$, where the a 's are distinct and alternate between 1 and 2. Obviously, $L(B_e) = n - 2$, and $L(B) \geq n$. If $B = a_1^2 a_2' a_3'' \cdots a_{n-1}' a_n$, then $L(B_e) = n - 1$, $L(B) \geq n + 1$. If finally $B = a_1^2 a_2^2 a_3'' \cdots a_{n-1}' a_n$, then $L(B_e) = n$, $L(B) \geq n + 2$. Thus (9) is valid. The validity of (10) is obvious.

Theorems 11-13 to follow will be needed in a later section.

THEOREM 11. Let T_{ee} and T_e be the exponent trajectories of trajectories T_e and T respectively, and suppose that T , T_e , and T_{ee} have the generating symbols 1, 2. Let JED be a subblock of T , and suppose that the blocks J , E , and D are so related that E is the proper exponent block of D , while J is the exponent block due to the join of E and D . If $L(D) \geq 4$, then

$$(11) \quad L(JE) < L(D).$$

We write $D = GH$, where G is a block of length 4. We let J_e denote the exponent block due to the join of G and H , and let G_e , H_e denote the proper exponent blocks of G and H respectively. We have the following relations:

$$(12) \quad L(JE) = L(J) + L(E),$$

$$(13) \quad L(E) = L(G_e) + L(J_e) + L(H_e).$$

We consider first the case where G begins with the block α^2 . By the assumption $L(G) = 4$ we have $G \neq \alpha$, α^2 , whence by Theorem 10 the relation $L(E) \leq L(D) - 2$ follows. Since D begins with α^2 , the block J contains no exponent arising from D . Hence J is vacuous or 1, whence $L(J) \leq 1$. It follows by (12) that (11) is valid.

Next, we suppose that G begins with $\alpha\beta$ ($\alpha \neq \beta$). If $G = \alpha\beta\alpha\alpha$, then $G_e = 12$. If H is vacuous, $E = 12$ and $ED = 12\alpha\beta\alpha\alpha$. If $\alpha = 1$, the proper exponent block of $12\alpha\beta\alpha\alpha$ contains a subblock 1^3 , which is impossible in view of the fact that T_{ee} contains only the symbols 1, 2. Hence $\alpha = 2$, and $J = 2$. Thus $L(JE) = 3$, and (11) holds. If H is not vacuous, the block GH begins with $\alpha\beta\alpha\alpha\beta$ since T_e contains only the generating symbols 1, 2. By Theorem 10, $L(E) \leq L(D) - 3$. Since $L(J) \leq 2$, formula (11) is valid. We now let $G = \alpha\beta\beta\alpha$. Since we have an intermediate block β^2 , by Theorem 10 we have $L(E) \leq L(D) - 3$, whence (11) holds. If finally $G = \alpha\beta\alpha\beta$, G is preceded in T by α since we cannot have a block 1^3 in T_e . Then $J = 2$, and $L(J) = 1$. By Theorem 10 we have $L(E) \leq L(D) - 2$, whence (11) holds. Thus in any case (11) is valid.

THEOREM 12. *Let T_e be the exponent trajectory of a trajectory T , and let T and T_e be trajectories in the generating symbols 1, 2. Let JED be a subblock of T , where J , E , and D are related as in Theorem 11. If $L(D) \geq 4$, then $L(JE) \geq 2$.*

If the leading 4-block of D is of the form $\alpha\alpha\beta\beta$, $\alpha\alpha\beta\alpha$, $\alpha\beta\alpha\alpha$, or $\alpha\beta\alpha\beta$ ($\alpha \neq \beta$), the proper exponent block of this block is of length 2, whence $L(JE) \geq 2$. If the leading 4-block of D is of the form $\alpha\beta\beta\alpha$, this block has the proper exponent block 2, whence $L(E) \geq 1$. The leading symbol α of D will yield an exponent in J . Thus in any case $L(JE) \geq 2$.

THEOREM 13. *Let T , T_e , J , E , and D be defined as in Theorem 12. If J is non-vacuous, then E is non-vacuous.*

6. Subrays of a trajectory identical with its exponent trajectory. The theorem which follows is valid for trajectories based on an arbitrarily given set of generating symbols, and is not restricted to the 1, 2 case.

THEOREM 14. *If a trajectory T is identical with its exponent trajectory T_e , the trajectory T does not contain two identical subrays R_1 , R_2 with initial elements in different positions in T .*

Suppose that the rays R_1 , R_2 are directed to the right in the sense that $R_1 = R_2 = abc \dots$. The rays R_1 and R_2 overlap, whence it is no restriction to suppose that R_1 overlaps R_2 . Let the subblock of R_1 which precedes R_2 in R_1 be denoted by B . Since $R_1 = R_2$, the ray R_2 contains a subray R_3 identical with R_2 and preceded in R_2 by the block B . Thus T contains the subray $N = BBB \dots$. Since $T = T_e$, the trajectory T_e contains a subray N_1 identical with the ray N . Let N_e denote the proper exponent ray of N . The rays N_1 and N_e overlap in T_e . Therefore the ray N_e contains a subray N_2 identical with N . Clearly, N_2 is the exponent ray of a subray $N_3 = B_1 B_1 B_1 \dots$ of N where B is the exponent block of B_1 . Since the ray N_3 is a subray of the ray N , and $l(B_1) \geq l(B)$, we can write B_1 as $B_{11} B^r B_{12}$, where $B_{12} B_{11} = B$, and $r \geq 0$. If $r = 0$ it is understood that the block B^r is vacuous. Thus the trajectory $T_1 = \dots B_1 B_1 B_1 \dots$ obtained by continuing N_3 to the left is identical with the trajectory $T_2 = \dots BBB \dots$. But T_2 is the exponent trajectory of T_1 , whence by Theorem 2 we have arrived at a contradiction.

7. The uniqueness of a trajectory identical with its exponent trajectory in the case of generating symbols 1, 2. We shall prove in this section that the trajectory (7) is the only one of its kind for trajectories in generating symbols 1, 2. We let T^{-1} denote the trajectory obtained from a trajectory T by reversing the order of the symbols in T .

LEMMA 3. *If a trajectory T is identical with its exponent trajectory T_e , and T*

contains the generating symbols 1, 2 only, the trajectory T or T^{-1} contains a subray

$$(14) \quad R = B'_1 B'_2 B'_3 \cdots,$$

where B'_i is the exponent block of B'_{i+1} , and the last symbol of B'_i is different from the first symbol of B'_{i+1} for each i .

Let a denote a symbol of T in a fixed position in T . The corresponding symbol a of the exponent trajectory T_e is the exponent of a symbol b in T so that the block b^a occurs in T . It is no restriction to assume that the block b^a is not to the left of the symbol a in T . We suppose first that the block b^a of T does not contain the symbol a , so that b^a is to the right of a in T .

We let B'_i denote the block of symbols in T starting with a and ending with the symbol preceding the block b^a in T . Since $T = T_e$, the symbol a in T_e is the initial symbol of a block B'_i in T_e . The block of T starting with b^a and having B'_i as exponent block is unique since consecutive exponents in T are exponents on distinct bases alternating between the symbols 1, 2. We emphasize that T is of the form

$$\cdots 1^{a-2} 2^{a-1} 1^{a0} 2^{a1} 1^{a2} \cdots$$

We denote the block of T starting with b^a and having exponent block B'_i by B'_i . Thus T contains the block $B'_1 B'_2$. We assume now that T contains the block $B'_1 B'_2 \cdots B'_r$ where $B'_1 B'_2 \cdots B'_{r-1}$ is the exponent block of B'_r . Since $T = T_e$, the block $B'_1 B'_2 \cdots B'_{r-1}$ in T_e is followed by B'_r , whence $B'_1 B'_2 \cdots B'_r$ in T is followed by a block B'_{r+1} whose exponent block is B'_r , and the first symbol of B'_{r+1} is distinct from the last symbol in B'_r . Thus T contains the subray R .

Finally, we suppose that the block b^a of T contains the symbol a of T . If $a=1$, then b^a is the block 1^1 . Since the bases alternate between 1 and 2, the block b^a is preceded and followed in T by the base 2. Thus T contains the block $B_0 = 2a2 = 212$, where B_0 is the block B_0 occurring in (7). Since $T = T_e$, the symbol a in T_e is preceded and followed by 2 in T_e , whence a is the central symbol in a block B_0 of T_e . It follows that B_0 is the exponent block of a block $B_1^{-1} B_0 B_1$ in T with central symbol a and $B_1 = 2$. Making use of the equality $T = T_e$ and developing T to the right and left of $B_1^{-1} B_0 B_1$ as in §4 we obtain (7). The subray

$$(15) \quad B_1 B_2 B_3 \cdots$$

of (7) is clearly a subray of the type (14). If now $a=2$, the symbol a in T is either the leading or final symbol in the block b^a , so that b^a is either $a2$ or $2a$. If $b^a = a2$, then since $a=2$, the block b^a is preceded in T by the symbol 1,

and thus this block is preceded in T_e by the symbol 1. Thus the block 1^1 precedes b^a in T , and since both base and exponent in 1^1 are followed by a in T and T_e respectively, the base and exponent in the power 1^1 are corresponding symbols. The argument thus reduces to the preceding case where $a=1$. If now $a=2$, while $b^a=2a$, the block b^a is followed in T_e by the symbol 1, so that the block b^a in T is followed by the power 1^1 . Clearly the base and exponent in this power are corresponding symbols, whence we have again reduced the argument to the case where $a=1$. Thus T contains the subray (14), and Lemma 3 is proved, for if b^a is to the left of a in T , then b^a is to the right of a in T^{-1} .

We remark that the exponent ray R_e of $B'_2 B'_3 B'_4 \dots$ in (14) is the ray R .

We consider now a trajectory T with $T=T_e$, whence by the lemma just proved T contains the subray R of (14). We let E_0 be a block such that the ray

$$(16) \quad E_0 B'_1 B'_2 \dots$$

is the proper exponent ray of R in (14). The block E_0 may be vacuous. Since $T=T_e$, the trajectory T contains (16) as a subray. We let J_0 denote the exponent block due to the join of E_0 and B'_1 in $E_0 B'_1$. It is clear that T contains the subray $J_0 E_0 B'_1 B'_2 \dots$. For $i>0$, we let E_i denote the proper exponent block of a block $J_{i-1} E_{i-1}$, and J_i the exponent block due to the join of E_i and $J_{i-1} E_{i-1}$. In the following lemma we use G_i to denote the block $J_i E_i \dots J_1 E_1 J_0 E_0$, and R as in (14). Here $G_{i+1}=G_i$ if $J_{i+1} E_{i+1}$ is vacuous.

LEMMA 4. *If a trajectory T is equal to its exponent trajectory T_e , and T contains the subray*

$$(17) \quad G_i R,$$

the trajectory T contains the subray

$$(18) \quad G_{i+1} R,$$

where $J_{i+1} E_{i+1}$ in G_{i+1} may be vacuous.

We assume that T contains the subray (17). The proper exponent ray of (17) is the ray

$$(19) \quad E_{i+1} G_i R.$$

Since $T=T_e$, the trajectory T contains the subray (19). Evidently the proper exponent ray of (19) contains the subray (18).

Theorems 11 and 12 yield at once the following lemma.

LEMMA 5. *If the subray (17) in a trajectory T with $T=T_e$ is continued to the left, one arrives at a block $J_e E_e$ of length 2 or 3, provided the trajectory T contains a subblock $J_i E_i$ of length at least 2.*

LEMMA 6. *If a trajectory T in generating symbols 1, 2 is identical with its exponent trajectory T_e , the trajectory T or T^{-1} contains a subray identical with the subray (15) of (7).*

By Lemma 3, the trajectory T or T^{-1} contains a subray (14). Suppose that T contains (14). We continue the subray (14) of T to the left to obtain a subray (17) of T where E_{i+1} is vacuous. We suppose first that the subray (17), which is explicitly the ray

$$J_i E_i \cdots J_1 E_1 J_0 E_0 B'_1 B'_2 B'_3 \cdots$$

contains a block $J_i E_i$ of length at least 2, whence by Lemma 5 the ray (17) contains a subblock $J_e E_e$ of length 2 or 3.

We assume that $L(J_e E_e) = 2$. If $J_e E_e = 22$, then $E_{\sigma+1} = 2$, and $E_{\sigma+1} J_e E_e$ contains the subblock 2^3 . Hence $J_e E_e \neq 22$. Suppose that $J_e E_e = 12$. We cannot have $J_e = 12$, since by Theorem 13 the block E_e is then not vacuous. Also, we cannot have $E_e = 12$ since the symbol 2 in E_e yields an exponent in J_e due to the join of E_e with the block following E_e in T , whence J_e is not vacuous. Thus $J_e = 1$, and $E_e = 2$. Then $J_e E_e$ is followed by a block 122 in T . Now $J_e E_e 122$ has the proper exponent block $112 = 1 J_e E_e$. Since $T = T_e$, the trajectory T contains the subray $R_e = 112 G_{e-1} R$, where if $\sigma = 0$, we understand that $G_{e-1} = G_{-1}$ is vacuous. The leading block 11 in R_e has exponent 2, whence T contains the subray $R_b = K G_e R$, where $K = 21$. The proper exponent ray of R_b is R_b itself. We write B_1 for the symbol 2 in K , and B_2 for the block 11 in $K J_e$. The exponent of B_2 in R_b is the initial symbol of the proper exponent ray of R_b . If we define B_3, B_4, \dots as in §4, it is clear that R_b is identical with the ray (15). If $J_e E_e = 11$, then $E_{\sigma+1} = 2$. Writing $B_1 = B_{\sigma+1}$, $B_2 = J_e E_e$, defining B_i ($i > 2$) as in §4, and using the fact that (19) is the proper exponent ray of (17), we find that in this case (17) with $i = \sigma + 1$ is identical with the subray (15) of (7). Finally, we write $J_e E_e = 21$. If $E_e = 21$, the symbol 1 in E_e yields an exponent so that J_e is not vacuous. Hence $J_e = 2$, $E_e = 1$. Now $J_e E_e$ is followed in T by the block 121. Writing B_1 for J_e , and B_2 for the block 11 which follows J_e in $J_e E_e 121$, and defining B_i ($i > 2$) as in §4, we find that in this case (17) with $i = \sigma$ is identical with (15).

If $J_e E_e = 221$, then $E_{\sigma+1} = 2$ and $E_{\sigma+1} J_e E_e = 2^3 1$ which is impossible. If $J_e E_e = 211$, then $E_{\sigma+1} = 2$, $J_{\sigma+1} = 2$, which in the paragraph above was proved impossible. If $J_e E_e = 212$, then $E_{\sigma+1} = J_{\sigma+1} = 1$, which case was treated above. If $J_e E_e = 121$, then $E_{\sigma+1} = 1$, $J_{\sigma+1} = 2$, which was also treated above. If $J_e E_e = 112$, then $E_{\sigma+1} = 2$, and $J_{\sigma+1}$ is vacuous. Writing $B_1 = E_{\sigma+1}$, and B_2 for the leading block 11 of $J_e E_e$, and defining B_i ($i > 2$) as in §4, it is clear that (17) with $i = \sigma + 1$ is in this case identical with (15). If $J_e E_e = 122$, then

$J_{\sigma+1}E_{\sigma+1}=12$, which case was treated above. This completes the cases where $L(J_{\sigma}E_{\sigma})=3$, since the blocks 111 and 222 cannot occur in T .

Suppose now that T contains no block $J_{\sigma}E_{\sigma}$ with $L(J_{\sigma}E_{\sigma}) \geq 2$. Assume that T contains a subblock $J_{\sigma}E_{\sigma}$ with $L(J_{\sigma}E_{\sigma})=1$. We cannot have $J_{\sigma}E_{\sigma}=1$, whence $E_{\sigma}=1$, since E_{σ} is the proper exponent block of 121 or 212, and the block $E_{\sigma}121$ or $E_{\sigma}212$ yields a non-vacuous block J_{σ} due to the join of E_{σ} with 121 or 212. If $J_{\sigma}E_{\sigma}=2$, then $E_{\sigma}=2$, and E_{σ} is followed by 11 in T . Writing $B_1=J_{\sigma}E_{\sigma}$, $B_2=11$, and defining B_i ($i>2$) as in §4, we find that T contains the subray (15).

We suppose, finally, that T contains no subray (17) with a block $J_{\sigma}E_{\sigma}$ for which $L(J_{\sigma}E_{\sigma}) \geq 1$. Thus the subray (14) of T cannot be continued to the left. The block B'_1 cannot be of length greater than or equal to 3 since then B'_1 would yield a non-vacuous block E_0 . In the same way $B'_1 \neq 11, 22$. If $B'_1=12$, then $B'_2=122$, and $E_0=1$, whereas if $B'_1=21$, then $B'_2=221$, and $E_0=1$. Thus $L(B'_1)=1$. If $B'_1=1$, then $B'_2=2$, $B'_3=11$. In this case dropping the first block in (14), we obtain the subray (15) of (7) as a subray by writing $B_i=B'_{i+1}$, $i \geq 1$. If $B'_1=2$, writing $B_i=B'_i$ we obtain (15) from (14).

Thus in any case the trajectory T or T^{-1} contains the subray (15) of (7).

THEOREM 15. *There is one and only one trajectory T in generating symbols 1, 2 identical with its exponent trajectory.*

By Theorem 14 and Lemma 6 the trajectory T or T^{-1} contains the subray (15) of the trajectory (7) exactly once. Suppose that (15) is a subray of T . If (15) is preceded by the symbol 2 in T , the subray $\rho_1=2B_1B_2 \cdots$ of T must be preceded by the symbol 1, since no block 2^3 can occur in T . Thus T contains the subray $\rho_2=12B_1B_2 \cdots$. The proper exponent ray of ρ_1 is ρ_1 itself. In particular since the proper exponent ray ρ_1 of ρ_2 is preceded by the symbol 1, the ray ρ_2 is preceded in T by the symbol 2, so that $212B_1B_2 \cdots$ occurs in T . We write $B_0=212$, whence the ray $\rho_3=B_0B_1B_2 \cdots$ occurs in T . Since ρ_3 occurs in T , the ray $\rho_4=2212B_1B_2 \cdots$ occurs in T , whence $B_1^{-1}B_0B_1B_2 \cdots$ occurs in T . By induction, since

$$B_{r-1}^{-1} \cdots B_1^{-1}B_0B_1B_2 \cdots$$

occurs in T , the same ray occurs in T , and is the exponent ray of the ray

$$B_r^{-1} \cdots B_1^{-1}B_0B_1B_2 \cdots$$

Thus T is identical with (7).

If, on the other hand, the subray (15) is preceded by 1 in T , the trajectory T contains either the subray $R_1=21B_1B_2 \cdots$ or the subray $R_2=11B_1B_2 \cdots$. The proper exponent rays of R_1 and R_2 are respectively R_2 and R_1 . Since

$T = T_*$, the trajectory T then contains both R_1 and R_2 . Since $R_1 \neq R_2$, the sub-ray (15) in R_1 and R_2 occurs twice in T , contradicting Theorem 14.

If T is the trajectory (7), then $T = T^{-1}$. Thus Theorem 15 is proved.

THEOREM 16. *The trajectory (7) is the exponent trajectory of two distinct trajectories in generating symbols 1, 2.*

Theorem 16 states that (7) is not symmetric in the symbols 1, 2. Suppose, on the contrary, that (7) is unchanged when we interchange the symbols 1 and 2. Let C_i be the block obtained from B_i by interchanging 1 and 2 in B_i . Then we have the trajectory

$$(20) \quad \dots C_2^{-1} C_1^{-1} C_0 C_1 C_2 \dots$$

We remark that the block C_i , $i \geq 2$, has the exponent block B_{i-1} , whence C_i^{-1} has the exponent block B_{i-1}^{-1} . The block $C_1^{-1} C_0 C_1$ has the exponent block B_0 . We note that the symbol 2 in $C_1^{-1} C_0 C_1 = 11211$ yields the exponent 1 in B_0 . Now the ray $C_1^{-1} C_0 C_1 C_2 \dots$ has the proper exponent ray $\sigma = B_0 B_1 B_2 \dots$. If the trajectory (20) is identical with the trajectory (7), the trajectory (20) contains a subray $\sigma' = B_1^{-1} \sigma$ with proper exponent ray σ , where the symbol 1 in the subblock B_0 of the exponent ray σ is the exponent of the symbol 1 in the subblock B_0 of the ray σ' . Thus the exponent trajectory (7) of (20) contains the subray σ twice with initial symbols of each σ in different positions in (7). By Theorem 14 we have arrived at a contradiction.

Although (7) is the exponent trajectory of two distinct trajectories T_1 and T_2 in generating symbols 1, 2, the trajectories T_1 and T_2 are equivalent in the sense that these trajectories differ only in the notation used for the generating symbols.

ARMOUR INSTITUTE OF TECHNOLOGY,
CHICAGO, ILL.

A CORRECTION TO "THE BOUNDARY PROBLEM OF AN ORDINARY LINEAR DIFFERENTIAL SYSTEM IN THE COMPLEX DOMAIN"*

BY

RUDOLPH E. LANGER

In formula (6.1) replace $x_*^{(h, l)}$ by $x_*^{(h, \nu)}$, and $\mathfrak{R}(\lambda)$ by $\mathfrak{R}_\nu(\lambda)$, and add $\nu = 1, 2, \dots, n$. To derive (6.3) (with the accidentally omitted sign of integration from $x_*^{(h, l)}$ to x over the respective members of the sum), multiply (6.1) by $\mathfrak{S}(x)$ on the left, by $\mathfrak{C}(\lambda)\mathfrak{F}_{\nu, \nu}\mathfrak{S}^{-1}(x)$ on the right, and sum as to ν . In this formula and everywhere subsequently replace $\mathfrak{R}(\lambda)\mathfrak{C}(\lambda)$ by $\sum_{\nu=1}^n \mathfrak{R}_\nu(\lambda)\mathfrak{C}(\lambda)\mathfrak{F}_{\nu, \nu}$. The argument given shows that each $\mathfrak{R}_\nu(\lambda)$ is nonsingular. In and just before (6.9) replace $\mathfrak{R}^{-1}(\lambda)$ by $\sum_{\nu=1}^n \mathfrak{R}_\nu^{-1}(\lambda)\mathfrak{F}_{\nu, \nu}$. The stated result follows. (Throughout the discussion the hitherto undefined points $x_*^{(h, l)}$ with $h=l$, and the paths from them may be chosen arbitrarily in X .)

* Received by the editors October 5, 1939. Cf. these Transactions, vol. 46 (1939), pp. 151-190.

UNIVERSITY OF WISCONSIN,
MADISON, WIS.

A CORRECTION TO "PROPERTIES OF FUNCTIONS $f(x, y)$ OF BOUNDED VARIATION"*

BY

C. RAYMOND ADAMS AND JAMES A. CLARKSON

Some time after our paper was written it came to our attention that the partial derivatives of a measurable function $f(x, y)$ need not be measurable, in contradiction to a Lemma of Burkill and Haslam-Jones.† Trivial examples suffice to show this; indeed such an example, due to Hahn, has been given by Neubauer.‡ The proof of Theorem 18 of our paper, which made use of this lemma, is therefore unsound. Whether or not this theorem and its Corollary 2 are true we have been unable to determine. Corollary 1, however, to the effect that a function $f(x, y)$ in class $\bar{T} \cdot M$ has an approximate total differential almost everywhere, whose proof was our main objective, can readily be established as follows. Since f is in M , by a theorem of Saks§ the approximate partial Dini derivatives (or derivative numbers) are measurable functions; since f is in \bar{T} , the approximate partial derivatives are then measurable functions and are finite almost everywhere. The approximate total differentiability of f may then be inferred from a theorem of Stepanoff.||

* Received by the editors October 21, 1939. Cf. these Transactions, vol. 36 (1934), pp. 711-730.

† *Notes on the differentiability of functions of two variables*, Journal of the London Mathematical Society, vol. 7 (1932), pp. 297-305, Lemma 2.

‡ *Über die partiellen Derivierten unstetiger Funktionen*, Monatshefte für Mathematik und Physik, vol. 38 (1931), pp. 139-146, §1.

§ Saks, *Théorie de l'Intégrale*, Warsaw, 1933, p. 226, Theorem 2.

|| See, for example, Saks, loc. cit., p. 228, Theorem 3.

BROWN UNIVERSITY,
PROVIDENCE, R. I.,
THE UNIVERSITY OF PENNSYLVANIA,
PHILADELPHIA, PA.

